(54) Title: REGULATION OF HUMAN MEMBRANE-TYPE SERINE PROTEASE

(57) Abstract: Reagents which regulate human membrane-type serine protease activity and reagents which bind to human membrane-type serine protease gene products can be used to regulate extracellular matrix degradation. Such regulation is particularly useful for treating cancer (e.g., suppressing metastasis of malignant cells), neurodegenerative diseases, osteoporosis, and chronic obstructive pulmonary disease.

(57) 要約:

　マグネット１２を有したロータ２と、このロータ２を回転駆動するための磁界を発生するコイル２２と、このコイル２２に外部から給電するための給電端子１７とを備える。この給電端子１７からコイル２２の端末に至る全ての電気接続部が溶接部３１ｈによって接合されている。

- 1 -

REGULATION OF HUMAN MEMBRANE-TYPE SERINE PROTEASE

TECHNICAL FIELD OF THE INVENTION

5

The invention relates to the area of regulation human membrane-type serine protease activity to provide therapeutic effects.

BACKGROUND OF THE INVENTION

10

Metastasizing cancer cells invade the extracellular matrix using plasma membrane protrusions that contact and dissolve the matrix with proteases. Agents which inhibit such protease activity can be used to suppress metastases. Proteases also are expressed during development, when degradation of the extracellular matrix is

15   desired. In cases where appropriate extracellular matrix degradation does not occur, supplying a molecule with a protease activity can provide the necessary enzymatic activity. Thus, there is a need in the art for identifying new proteases and methods of regulating extracellular matrix degradation.

20   ### SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating human membrane-type serine protease. These and other objects of the invention are provided by one or more of the embodiments described below.

25

One embodiment of the invention is a membrane-type serine protease polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid

30   sequence shown in SEQ ID NO: 2; and

- 2 -

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7; and

5

the amino acid sequence shown in SEQ ID NO: 7.

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a

10　　membrane-type serine protease polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2;

15

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7; and

20

the amino acid sequence shown in SEQ ID NO: 7.

Binding between the test compound and the membrane-type serine protease polypeptide is detected. A test compound which binds to the membrane-type serine

25　　protease polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the membrane-type serine protease.

Another embodiment of the invention is a method of screening for agents which

30　　decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a membrane-type serine protease polypeptide, wherein the

- 3 -

polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 6; and

the nucleotide sequence shown in SEQ ID NO: 6.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the membrane-type serine protease through interacting with the membrane-type serine protease mRNA.

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a membrane-type serine protease polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7; and

- 4 -

the amino acid sequence shown in SEQ ID NO: 7.

A membrane-type serine protease activity of the polypeptide is detected. A test compound which increases membrane-type serine protease activity of the polypeptide relative to membrane-type serine protease activity in the absence of the test compound is thereby identified as a potential agent for increasing extracellular matrix degradation. A test compound which decreases membrane-type serine protease activity of the polypeptide relative to membrane-type serine protease activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a membrane-type serine protease product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 6; and

the nucleotide sequence shown in SEQ ID NO: 6.

Binding of the test compound to the membrane-type serine protease product is detected. A test compound which binds to the membrane-type serine protease product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

- 5 -

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a membrane-type serine protease polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 6; and

the nucleotide sequence shown in SEQ ID NO: 6.

Membrane-type serine protease activity in the cell is thereby decreased.

The invention thus provides reagents and methods for regulating human membrane-type serine protease activity that can be used *inter alia*, to suppress metastatic activity of malignant cells and to enhance extracellular matrix degradation during development.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the DNA-sequence encoding a membrane-type serine protease polypeptide (SEQ ID NO: 1).

Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1 (SEQ ID NO: 2).

Fig. 3 shows the amino acid sequence of the protein identified by Swiss Prot Accession No. P56677 (SEQ ID NO:3).

Fig. 4 shows the amino acid sequence of a membrane-type serine protease polypeptide (SEQ ID NO: 4).

Fig. 5 shows the amino acid sequence of a membrane-type serine protease polypeptide (SEQ ID NO:5).

Fig. 6 shows the DNA-sequence encoding a membrane-type serine protease polypeptide (SEQ ID NO: 6)

Fig. 7 shows the the amino acid sequence deduced from the DNA-sequence of Fig. 6 (SEQ ID NO: 7).

Fig. 8 shows the BLASTP alignment of SEQ ID NO:2 with SEQ ID NO:3 P56677).

Fig. 9 shows the BLOCKS search results.

Fig. 10 shows the HMMPFAM alignment of SEQ ID NO:2 with pfam|hmm|trypsin (SEQ ID NO:14).

Fig. 11 shows the HMMPFAM alignment of SEQ ID NO:2 with pfam|hmm|ldl_recept_a (SEQ ID NO:15).

Fig. 12 shows the BLASTP alignment of SEQ ID NO:7 with SEQ ID NO:3 (P56677).

Fig. 13 shows the relative expression of human membrane-type serine protease in various human tissues.

Fig. 14 shows the relative expression of human membrane-type serine protease in various human respiratory tissues and cells.

Fig. 15 shows the prosite search results.

Fig. 16 shows the BLOCKS search results.

Fig. 17 shows the HMMPFAM - alignment of 154_v4_TR1 against pfam|hmm|trypsin.

Fig. 18 shows the HMMPFAM - alignment of 154_v4_TR1 (SEQ ID NO:6) against pfam|hmm|ldl_recept_a.

Fig. 19 shows the HMMPFAM - alignment of 154_v4_TR1 (SEQ ID NO:6) against pfam|hmm|ldl_recept_a.

Fig. 20 shows the HMMPFAM - alignment of 154_v4_TR1 (SEQ ID NO:6) against pfam|hmm|ldl_recept_a.

- 7 -

Fig. 21 shows the exon-intron structure of the human membrane-type serine protease (154_v4): Two refseq genomic sequences were used: NT_022484.2, annotated to be on chromosome 3 3q13.2-3q13.33 and NT_023818.2 annotated to be on chromosome 8 with location undetermined. The NT_023818.2 covers about half of NT_022484.2, with almost identical sequence but a few nucleotide difference. A combination of Genscan, Genewise and Geneid was used to predict the exons. Experimental data was used to determine the validity of some exons and resolve the discrepancy between the available genomic sequences. Exon/intron structure in alignment with NT_022484.2 is shown. Alignment with NT_023818.2 is provided for the exon with discrepancy to NT_022484.2.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide encoding a membrane-type serine protease polypeptide and being selected from the group consisting of:

a)   a polynucleotide encoding a membrane-type serine protease polypeptide comprising an amino acid sequence selected from the group consisting of: amino acid sequences which are at least about 50% identical to

the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7; and

the amino acid sequence shown in SEQ ID NO: 7.

b)   a polynucleotide comprising the sequence of SEQ ID NOS: 1 or 6;

c)   a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

d)   a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

- 8 -

e)   a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d).

Furthermore, it has been discovered by the present applicant that a new membrane-type serine protease, particularaly a human membrane-type serine protease, has been identified. BLAST alignments show very good similarity with epithin, as well as with other cell surface serine proteases, such as mast cell protease 6. All active site residues for serine protease are present. Fibronectin domains also are present, indicating a cell surface association. The protein also contains other typical protease domains, such as Kringle and apple domains.

The C-terminal portion of human membrane-type serine protease as shown in SEQ ID NO:2 is 43% identical over 344 amino acids to the mouse protein identified by SwissProt Accession No. P56677 (SEQ ID NO:3) and annotated as mouse epithin (FIG. 1). The putative full-length human membrane-type serine protease (SEQ ID NO:7) is 30% identical over 862 amino acids to SEQ ID NO:3 (FIG. 5). A similar mouse homolog (likely an ortholog) is expressed in cerebellum and corpora quadri-gemina.

Modulation of the activity of human membrane-type serine protease is therefore expected to be useful for the same purposes as previously identified serine proteases, particularly in connection with the degradation of extracellular matrix and for treating diseases such as cancer, osteoporosis, chronic obstructive pulmonary disease (COPD), and various CNS diseases.

Polypeptides

Membrane-type serine protease polypeptides according to the invention comprise at least 6, 10, 15, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, or 350 contiguous amino acids selected from SEQ ID NO:6 or from a biologically active variant thereof, as defined below. A membrane-type serine protease polypeptide of

- 9 -

the invention therefore can be a portion of a membrane-type serine protease molecule, a full-length membrane-type serine protease molecule, or a fusion protein comprising all or a portion of a membrane-type serine protease molecule.

Biologically Active Variants

Membrane-type serine protease variants which are biologically active, *i.e.*, retain a membrane-type serine protease activity, also are membrane-type serine protease polypeptides. Preferably, naturally or non-naturally occurring membrane-type serine protease variants have amino acid sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to an amino acid sequence shown in SEQ ID NOs:2 or 7. Percent identity between a putative membrane-type serine protease variant and an amino acid sequence of SEQ ID NOS:2 or 7 is determined with the Needleman/Wunsch algorithm (Needleman and Wunsch, J.Mol. Biol. 48; 443-453, 1970) using a Blosum62 matrix with a gap creation penalty of 8 and a gap extension penalty of 2 (S. Henikoff and J.G. Henikoff, Proc. Natl. Acad. Sci. USA 89:10915-10919, 1992).

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active membrane-type serine protease polypeptide can

- 10 -

readily be determined by assaying for fibronectin binding or for membrane-type serine protease activity, as is known in the art and described, for example, in Example 4.

Fusion Proteins

Fusion proteins are useful for generating antibodies against membrane-type serine protease amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins which interact with portions of a membrane-type serine protease polypeptide, including its active site and fibronectin domains. Methods such as protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A membrane-type serine protease fusion protein comprises two protein segments fused together by means of a peptide bond. For example, the first protein segment can comprise at least 6, 10, 15, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, or 350 contiguous amino acids of SEQ ID NOS:2 or 7 or a biologically active variant thereof. Preferably, a fusion protein comprises the active site of the protease and/or one or more of the functional domains identified in FIGS. 1-5. The first protein segment also can comprise full-length membrane-type serine protease.

The second protein segment can be a full-length protein or a protein fragment or polypeptide. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose

- 11 -

binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the membrane-type serine protease polypeptide-encoding sequence and the heterologous protein sequence, so that the membrane-type serine protease polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two protein segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises membrane-type serine protease coding sequences disclosed herein in proper reading frame with nucleotides encoding the second protein segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologs

Species homologs of human membrane-type serine protease can be obtained using membrane-type serine protease polynucleotides (described below) to make suitable probes or primers to screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of membrane-type serine protease, and expressing the cDNAs as is known in the art.

Polynucleotides

A membrane-type serine protease polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a membrane-type serine protease polypeptide. A coding sequence for the C-terminal portion of membrane-type serine protease (SEQ ID NO:2) is shown in SEQ ID NO:1. A coding sequence for full-length membrane-type serine protease (SEQ ID NO:7) is shown in SEQ ID NO:6.

Degenerate nucleotide sequences encoding human membrane-type serine protease polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to the membrane-type serine protease coding sequence shown in SEQ ID NO:1 or 6 also are membrane-type serine protease polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of membrane-type serine protease polynucleotides which encode biologically active membrane-type serine protease polypeptides also are membrane-type serine protease polynucleotides.

Identification of Variants and Homologs

Variants and homologs of the membrane-type serine protease polynucleotides disclosed above also are membrane-type serine protease polynucleotides. Typically, homologous membrane-type serine protease polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known membrane-type serine protease polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50 °C once, 30 minutes; then 2X SSC, room temperature twice, 10

- 13 -

minutes each--homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the membrane-type serine protease polynucleotides disclosed herein can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of membrane-type serine protease polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5 °C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol. 81*, 123 (1973). Variants of human membrane-type serine protease polynucleotides or membrane-type serine protease polynucleotides of other species can therefore be identified, for example, by hybridizing a putative homologous membrane-type serine protease polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1 or 6 to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising membrane-type serine protease polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to membrane-type serine protease polynucleotides or their complements following stringent hybridization and/or wash conditions are also membrane-type serine protease polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20 °C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a membrane-

- 14 -

type serine protease polynucleotide having a coding sequence disclosed herein and a polynucleotide sequence which is at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to that nucleotide sequence can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A.*

5   *48*, 1390 (1962):

$T_m = 81.5 \,°C - 16.6(\log_{10} [Na^+]) + 0.41(\%G + C) - 0.63(\%formamide) - 600/l)$,

where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65 °C, or 50%

10   formamide, 4X SSC at 42 °C, or 0.5X SSC, 0.1% SDS at 65 °C. Highly stringent wash conditions include, for example, 0.2X SSC at 65 °C.

Preparation of Polynucleotides

15   A naturally occurring membrane-type serine protease polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or synthesized using an automatic synthesizer.

20   Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated membrane-type serine protease polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments which comprise membrane-type serine protease nucleotide sequences. Isolated polynucleotides are in preparations

25   which are free or at least 70, 80, or 90% free of other molecules.

Membrane-type serine protease cDNA molecules can be made with standard molecular biology techniques, using membrane-type serine protease mRNA as a template. Membrane-type serine protease cDNA molecules can thereafter be

30   replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR,

- 15 -

can be used to obtain additional copies of membrane-type serine protease polynucleotides, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesize membrane-type serine protease polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a membrane-type serine protease polypeptide having, for example, the amino acid sequence shown in SEQ ID NOS:2 or 7 or a biologically active variant of that sequence.

Obtaining Full-Length Polynucleotides

Various PCR-based methods can be used to extend the nucleic acid sequences encoding the disclosed portions of human membrane-type serine protease to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic. 2*, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al., Nucleic Acids Res. 16*, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68 - 72 °C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

- 16 -

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic. 1*, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations are used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res. 19*, 3055-3060, 1991. Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA. This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

- 17 -

Obtaining Polypeptides

Membrane-type serine protease polypeptides can be obtained, for example, by purification from human cells, by expression of membrane-type serine protease polynucleotides, or by direct chemical synthesis.

Protein Purification

Membrane-type serine protease polypeptides can be purified from cells, such as primary tumor cells, metastatic cells, or cancer cell lines (*e.g.*, colon cancer cell lines HCT116, DLD1, HT29, Caco2, SW837, SW480, and RKO, breast cancer cell lines 21-PT, 21-MT, MDA-468, SK-BR3, and BT-474, the A549 lung cancer cell line, or the H392 glioblastoma cell line), as well as cells transfected with a membrane-type serine protease expression construct. Brain tissues, such as cerebellum and corpora quadrigemina, also are useful sources. A purified membrane-type serine protease polypeptide is separated from other compounds which normally associate with the membrane-type serine protease polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified membrane-type serine protease polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis. Enzymatic activity of the purified preparations can be assayed, for example, as described in Example 4.

Expression of Polynucleotides

To express a membrane-type serine protease polypeptide, a membrane-type serine protease polynucleotide can be inserted into an expression vector which contains the

- 18 -

necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding membrane-type serine protease polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y, 1989.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a membrane-type serine protease polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In

- 19 -

mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a membrane-type serine protease polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the membrane-type serine protease polypeptide. For example, when a large quantity of a membrane-type serine protease polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the membrane-type serine protease polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem. 264*, 5503-5509, 1989 or pGEX vectors (Promega, Madison, Wis.) can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or Factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al., Methods Enzymol. 153*, 516-544, 1987.

- 20 -

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding membrane-type serine protease polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV (Takamatsu *EMBO J. 6*, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al., EMBO J. 3*, 1671-1680, 1984; Broglie *et al., Science 224*, 838-843, 1984; Winter *et al., Results Probl. Cell Differ. 17*, 85-105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs or Murray, in MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

An insect system also can be used to express a membrane-type serine protease polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding membrane-type serine protease polypeptides can be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of membrane-type serine protease polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect, for example, *S. frugiperda cells* or *Trichoplusia* larvae in which membrane-type serine protease polypeptides can be expressed (Engelhard *et al., Proc. Nat. Acad. Sci. 91*, 3224-3227, 1994).

Mammalian Expression Systems

A number of viral-based expression systems can be utilized in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding membrane-type serine protease polypeptides can be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing a membrane-type serine protease polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci.* *81*, 3655-3659, 1984). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding membrane-type serine protease polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a membrane-type serine protease polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell

- 22 -

system which is used (see Scharf *et al.*, *Results Probl. Cell Differ.* 20, 125-162, 1994).

Host Cells

5

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process an expressed membrane-type serine protease polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation,

10 lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection

15 (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express membrane-type serine

20 protease polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the

25 selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced membrane-type serine protease sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

30 Any number of selection systems can be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase

- 23 -

(Wigler *et al.*, *Cell 11*, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell 22*, 817-23, 1980). Genes which can be employed in *tk⁻* or *aprt⁻* cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci.* 77, 3567-70, 1980); *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol. 150*, 1-14, 1981); and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992 *supra*). Additional selectable genes have been described, for example *trpB*, which allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci.* 85, 8047-51, 1988). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol.* 55, 121-131, 1995).

Detecting Expression of Polypeptides

Although the presence of marker gene expression suggests that the membrane-type serine protease polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a membrane-type serine protease polypeptide is inserted within a marker gene sequence, transformed cells containing sequences which encode a membrane-type serine protease polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a membrane-type serine protease polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the membrane-type serine protease polynucleotide.

Alternatively, host cells which contain a membrane-type serine protease polynucleotide and which express a membrane-type serine protease polypeptide can

- 24 -

be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or

5   protein.

The presence of a polynucleotide sequence encoding a membrane-type serine protease polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a

10   membrane-type serine protease polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a membrane-type serine protease polypeptide to detect transformants which contain a membrane-type serine protease polynucleotide.

15   A variety of protocols for detecting and measuring the expression of a membrane-type serine protease polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal

20   antibodies reactive to two non-interfering epitopes on a membrane-type serine protease polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med. 158*, 1211-1216, 1983).

25

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding membrane-type serine protease polypeptides include

30   oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a membrane-type serine protease

- 25 -

polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase, such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a membrane-type serine protease polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode membrane-type serine protease polypeptides can be designed to contain signal sequences which direct secretion of membrane-type serine protease polypeptides through a prokaryotic or eukaryotic cell membrane.

Other constructions can be used to join a sequence encoding a membrane-type serine protease polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the

membrane-type serine protease polypeptide can be used to facilitate purification.
One such expression vector provides for expression of a fusion protein containing a
membrane-type serine protease polypeptide and 6 histidine residues preceding a
thioredoxin or an enterokinase cleavage site. The histidine residues facilitate
5   purification on IMAC (immobilized metal ion affinity chromatography as described
in Porath *et al., Prot. Exp. Purif. 3*, 263-281, 1992), while the enterokinase cleavage
site provides a means for purifying the membrane-type serine protease polypeptide
from the fusion protein. Vectors which contain fusion proteins are disclosed in Kroll
et al., DNA Cell Biol. 12, 441-453, 1993).

10

Chemical Synthesis

Sequences encoding a membrane-type serine protease polypeptide can be
synthesized, in whole or in part, using chemical methods well known in the art (see
15   Caruthers *et al., Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al. Nucl. Acids
Res. Symp. Ser.* 225-232, 1980). Alternatively, a membrane-type serine protease
polypeptide itself can be produced using chemical methods to synthesize its amino
acid sequence. For example, membrane-type serine protease polypeptides can be
produced by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am.
20   Chem. Soc. 85*, 2149-2154, 1963; Roberge *et al., Science 269*, 202-204, 1995).
Protein synthesis can be performed using manual techniques or by automation.
Automated synthesis can be achieved, for example, using Applied Biosystems 431A
Peptide Synthesizer (Perkin Elmer). Various fragments of membrane-type serine
protease polypeptides can be separately synthesized and combined using chemical
25   methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high
performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND
MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The
30   composition of a synthetic membrane-type serine protease polypeptide can be
confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation

- 27 -

procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the membrane-type serine protease polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce membrane-type serine protease polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter membrane-type serine protease polypeptide-encoding sequences for a variety of reasons, including modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a membrane-type serine protease polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, F(ab')$_2$, and Fv, which are capable of binding an epitope of a membrane-type serine protease polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are

required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, e.g., at least 15, 25, or 50 amino acids.

5   An antibody which specifically binds to an epitope of a membrane-type serine protease polypeptide can be used therapeutically, as well as in immunochemical assays, including but not limited to Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or

10   immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to a membrane-type serine protease

15   polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to membrane-type serine protease polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a membrane-type serine protease polypeptide from solution.

20

Membrane-type serine protease polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a membrane-type serine protease polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and

25   keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (e.g., aluminum hydroxide), and surface active substances (e.g. lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used

30   in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially useful.

- 29 -

Monoclonal antibodies which specifically bind to a membrane-type serine protease polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature 256*, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods 81*, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci. 80*, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol. 62*, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci. 81*, 6851-6855, 1984; Neuberger *et al.*, *Nature 312*, 604-608, 1984; Takeda *et al.*, *Nature 314*, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, one can produce humanized antibodies using recombinant methods, as described in GB2188638B. Antibodies which specifically bind to a membrane-type serine protease polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies which specifically bind to membrane-type serine protease polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain

- 30 -

shuffling from random combinatorial immunoglobin libraries (Burton, *Proc. Natl. Acad. Sci. 88*, 11120-23, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev. 5*, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol. 15*, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem. 269*, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology. Verhaar *et al.*, 1995, *Int. J. Cancer 61*, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth. 165*, 81-91.

Antibodies which specifically bind to membrane-type serine protease polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci. 86*, 3833-3837, 1989; Winter *et al.*, *Nature 349*, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

- 31 -

Antibodies of the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a membrane-type serine protease polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of membrane-type serine protease gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. *See* Brown, *Meth. Mol. Biol. 20*, 1-8, 1994; Sonveaux, *Meth. Mol. Biol. 26*, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev. 90*, 543-583, 1990.

Modifications of membrane-type serine protease gene expression can be obtained by designing antisense oligonucleotides which will form duplexes to the control, 5', or regulatory regions of the membrane-type serine protease gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10

- 32 -

from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (e.g., Gee et al., in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful duplex formation between an antisense oligonucleotide and the complementary sequence of a membrane-type serine protease polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a membrane-type serine protease polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent membrane-type serine protease nucleotides, can provide targeting specificity for membrane-type serine protease mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular membrane-type serine protease polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a membrane-type serine protease polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5'

- 33 -

phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. *See, e.g.,* Agrawal *et al., Trends Biotechnol. 10,* 152-158, 1992; Uhlmann *et al., Chem. Rev. 90,* 543-584, 1990; Uhlmann *et al., Tetrahedron. Lett. 215,* 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. *See, e.g.,* Cech, *Science 236,* 1532-1539; 1987; Cech, *Ann. Rev. Biochem. 59,* 543-568; 1990, Cech, *Curr. Opin. Struct. Biol. 2,* 605-609; 1992, Couture & Stinchcomb, *Trends Genet. 12,* 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.,* Haseloff *et al.,* U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a membrane-type serine protease polynucleotide can be used to generate ribozymes which will specifically bind to mRNA transcribed from the membrane-type serine protease polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (*see* Haseloff *et al. Nature 334,* 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.,* EP 321,201).

- 34 -

Specific ribozyme cleavage sites within a membrane-type serine protease RNA target are initially identified by scanning the RNA molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the membrane-type serine protease target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. The suitability of candidate targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related; thus, upon hybridizing to the membrane-type serine protease target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease membrane-type serine protease expression. Alternatively, if it is desired that the cells stably retain the DNA construct, it can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. The DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors which induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of membrane-type serine protease mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

- 35 -

Differentially Expressed Genes

Described herein are methods for the identification of genes whose products interact with human membrane-type serine protease. Such genes may represent genes that are differentially expressed in disorders including, but not limited to, cancer, neuro-degenerative diseases, osteoporosis, and chronic obstructive pulmonary disease. Further, such genes may represent genes that are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human membrane-type serine protease gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

- 36 -

Identification of Differentially Expressed Genes

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique that does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples that represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature 308*, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 2825, 1984), and differential display (Liang & Pardee, *Science 257*, 967-71, 1992; U.S. Patent 5,262,311), and microarrays.

The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human membrane-type serine protease. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human membrane-type serine protease. The differential expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human membrane-type serine protease gene or gene product are up-regulated or down-regulated.

Screening Methods

The invention provides methods for identifying modulators, *i.e.*, candidate or test compounds which bind to membrane-type serine protease polypeptides or polynucleotides and/or have a stimulatory or inhibitory effect on, for example, expression or activity of the membrane-type serine protease polypeptide or polynucleotide, so as to regulate degradation of the extracellular matrix. Decreased extracellular matrix degradation is useful for preventing or suppressing malignant cells from metastasizing. Increased extracellular matrix degradation may be desired, for example, in developmental disorders characterized by inappropriately low levels of extracellular matrix degradation or in regeneration.

The invention provides assays for screening test compounds which bind to or modulate the activity of a membrane-type serine protease polypeptide or a membrane-type serine protease polynucleotide. A test compound preferably binds to a membrane-type serine protease polypeptide or polynucleotide. More preferably, a test compound decreases a membrane-type serine protease activity of a membrane-type serine protease polypeptide or expression of a membrane-type serine protease polynucleotide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods

- 38 -

requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. *See* Lam, *Anticancer Drug Des. 12*, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (*see*, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6909, 1993; Erb *et al. Proc. Natl. Acad. Sci. U.S.A. 91*, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem. 37*, 2678, 1994; Cho *et al.*, *Science 261*, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2061; Gallop *et al.*, *J. Med. Chem. 37*, 1233, 1994). Libraries of compounds can be presented in solution (*see*, *e.g.*, Houghten, *BioTechniques 13*, 412-421, 1992), or on beads (Lam, *Nature 354*, 82-84, 1991), chips (Fodor, *Nature 364*, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 89*, 1865-1869, 1992), or phage (Scott & Smith, *Science 249*, 386-390, 1990; Devlin, *Science 249*, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci. 97*, 6378-6382, 1990; Felici, *J. Mol. Biol. 222*, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

Test compounds can be screened for the ability to bind to membrane-type serine protease polypeptides or polynucleotides or to affect membrane-type serine protease activity or membrane-type serine protease gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

- 39 -

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 19*, 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity 2*, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

- 40 -

Binding Assays

For binding assays, the test compound is preferably a small molecule which binds to
and occupies the active site or a fibronectin domain of the membrane-type serine
protease polypeptide, thereby making the active site or fibronectin domain
inaccessible to substrate such that normal biological activity is prevented. Examples
of such small molecules include, but are not limited to, small peptides or peptide-like
molecules. In binding assays, either the test compound or the membrane-type serine
protease polypeptide can comprise a detectable label, such as a fluorescent,
radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase,
alkaline phosphatase, or luciferase. Detection of a test compound which is bound to
the membrane-type serine protease polypeptide can then be accomplished, for ex-
ample, by direct counting of radioemmission, by scintillation counting, or by
determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a membrane-type serine protease poly-
peptide can be determined without labeling either of the interactants. For example, a
microphysiometer can be used to detect binding of a test compound with a target
polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument
that measures the rate at which a cell acidifies its environment using a
light-addressable potentiometric sensor (LAPS). Changes in this acidification rate
can be used as an indicator of the interaction between a test compound and a
membrane-type serine protease polypeptide. (McConnell *et al.*, *Science 257*,
1906-1912, 1992).

Determining the ability of a test compound to bind to a membrane-type serine
protease polypeptide also can be accomplished using a technology such as real-time
Bimolecular Interaction Analysis (BIA). Sjolander & Urbaniczky, *Anal. Chem. 63*,
2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol. 5*, 699-705, 1995. BIA
is a technology for studying biospecific interactions in real time, without labeling any
of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface

- 41 -

plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a membrane-type serine protease polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell 72*, 223-232, 1993; Madura *et al.*, *J. Biol. Chem. 268*, 12046-12054, 1993; Bartel *et al.*, *BioTechniques 14*, 920-924, 1993; Iwabuchi *et al.*, *Oncogene 8*, 1693-1696, 1993; and Brent WO94/10300), to identify other proteins which bind to or interact with the membrane-type serine protease polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct a polynucleotide encoding a membrane-type serine protease polypeptide is fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence that encodes an unidentified protein ("prey" or "sample") is fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the membrane-type serine protease polypeptide.

It may be desirable to immobilize either the membrane-type serine protease polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommo-

- 42 -

date automation of the assay. Thus, either the membrane-type serine protease poly-peptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the membrane-type serine protease polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a membrane-type serine protease polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, a membrane-type serine protease polypeptide is a fusion protein comprising a domain that allows the membrane-type serine protease polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed membrane-type serine protease polypeptide; the mixture is then incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing polypeptides or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a membrane-type serine protease polypeptide (or polynucleotide) or a test compound

can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated membrane-type serine protease polypeptides or test compounds can be prepared from biotin-NHS(N-hydroxysuccinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a membrane-type serine protease polypeptide polynucleotides, or a test compound, but which do not interfere with a desired binding site, such as the active site or a fibronectin domain of the membrane-type serine protease polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the membrane-type serine protease polypeptide (or polynucleotides) or test compound, enzyme-linked assays which rely on detecting a membrane-type serine protease activity of the membrane-type serine protease polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a membrane-type serine protease polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a membrane-type serine protease polynucleotide or polypeptide can be used in a cell-based assay system. A membrane-type serine protease polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, including neoplastic cell lines such as the colon cancer cell lines HCT116, DLD1, HT29, Caco2, SW837, SW480, and RKO, breast cancer cell lines 21-PT, 21-MT, MDA-468, SK-BR3, and BT-474, the A549 lung cancer cell line, and the H392 glioblastoma cell line, can be used. An intact cell is contacted with a test compound. Binding of the test compound to a membrane-type serine protease polypeptide or polynucleotide is determined as described above, after lysing the cell to release the membrane-type serine protease polypeptide-test compound complex.

- 44 -

Enzyme Assays

Test compounds can be tested for the ability to increase or decrease a membrane-type serine protease activity of a membrane-type serine protease polypeptide. Membrane-type serine protease activity can be measured, for example, using the method described in Example 4. Membrane-type serine protease activity can be measured after contacting either a purified membrane-type serine protease polypeptide, a cell extract, or an intact cell with a test compound. A test compound which decreases membrane-type serine protease activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for decreasing extracellular matrix degradation. A test compound which increases membrane-type serine protease activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for increasing extracellular matrix degradation.

Gene Expression

In another embodiment, test compounds which increase or decrease membrane-type serine protease gene expression are identified. A membrane-type serine protease polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the membrane-type serine protease polynucleotide is determined. The level of expression of membrane-type serine protease mRNA or polypeptide in the presence of the test compound is compared to the level of expression of membrane-type serine protease mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of membrane-type serine protease mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of membrane-type serine protease mRNA or polypeptide is less expression. Alternatively, when expression of the mRNA or protein is less in the presence of the test compound

- 45 -

than in its absence, the test compound is identified as an inhibitor of membrane-type serine protease mRNA or polypeptide expression.

The level of membrane-type serine protease mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or protein. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a membrane-type serine protease polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a membrane-type serine protease polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses a membrane-type serine protease polynucleotide can be used in a cell-based assay system. The membrane-type serine protease polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, including neoplastic cell lines such as the colon cancer cell lines HCT116, DLD1, HT29, Caco2, SW837, SW480, and RKO, breast cancer cell lines 21-PT, 21-MT, MDA-468, SK-BR3, and BT-474, the A549 lung cancer cell line, and the H392 glioblastoma cell line, can be used.

Identification of Target and Pathway Genes and Proteins

Described herein are methods for the identification of genes whose products interact with human membrane-type serine protease. Such genes may represent genes which are differentially expressed in cancer, CNS disorders, COPD, and osteoporosis. Further, such genes may represent genes which are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Such differentially expressed genes may represent "target" and/or "fingerprint"

genes. Methods for the identification of such differentially expressed genes are described below. Methods for the further characterization of such differentially expressed genes, and for their identification as target and/or fingerprint genes also are described below.

In addition, methods are described for the identification of genes, termed "pathway genes," which are involved in cancer, CNS disorders, COPD, and osteoporosis. "Pathway gene," as used herein, refers to a gene whose gene product exhibits the ability to interact with gene products involved in these disorders. A pathway gene may be differentially expressed and, therefore, may have the characteristics of a target and/or fingerprint gene.

"Differential expression" refers to both quantitative as well as qualitative differences in a gene's temporal and/or tissue expression pattern. Thus, a differentially expressed gene may qualitatively have its expression activated or completely inactivated in normal versus diseased states, or under control versus experimental conditions. Such a qualitatively regulated gene will exhibit an expression pattern within a given tissue or cell type which is detectable in either normal or diseased subjects, but is not detectable in both. Alternatively, such a qualitatively regulated gene will exhibit an expression pattern within a given tissue or cell type which is detectable in either control or experimental subjects, but is not detectable in both. "Detectable" refers to an RNA expression pattern which is detectable via the standard techniques of differential display, RT-PCR and/or Northern analyses, which are well known to those of skill in the art.

A differentially expressed gene may have its expression modulated, *i.e.*, quantitatively increased or decreased, in normal versus diseased states, or under control versus experimental conditions. The degree to which expression differs in normal versus body weight disorder or control versus experimental states need only be large enough to be visualized via standard characterization techniques, such as, for example, the differential display technique described below. Other such standard

- 47 -

characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase) PCR and Northern analyses.

5   Differentially expressed genes may be further described as target genes and/or fingerprint genes. "Fingerprint gene" refers to a differentially expressed gene whose expression pattern may be utilized as part of a prognostic or diagnostic evaluation, or which, alternatively, may be used in methods for identifying compounds useful for the treatment of various disorders. A fingerprint gene may also have the character-

10   istics of a target gene or a pathway gene.

"Target gene" refers to a differentially expressed gene involved in cancer, COPD, CNS disorders, or osteoporosis by which modulation of the level of target gene expression or of target gene product activity may act to ameliorate symptoms. A

15   target gene may also have the characteristics of a fingerprint gene and/or a pathway gene.

- 48 -

Identification of Differentially Expressed Genes

A variety of methods may be utilized for the identification of genes which are involved in cancer, COPD, CNS disorders, or osteoporosis. To identify differentially expressed genes, RNA, either total or mRNA, may be isolated from one or more tissues of the subjects utilized in paradigms such as those described above. RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique which does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples which represent RNA produced by differentially expressed genes may be identified by utilizing a variety of methods which are well known to those of skill in the art. For example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature 308*, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 2825, 1984), and, preferably, differential display (Liang & Pardee, *Science 257*, 967-71, 1992; U.S. Patent 5,262,311), may be utilized to identify nucleic acid sequences derived from genes that are differentially expressed.

Differential screening involves the duplicate screening of a cDNA library in which one copy of the library is screened with a total cell cDNA probe corresponding to the mRNA population of one cell type while a duplicate copy of the cDNA library is screened with a total cDNA probe corresponding to the mRNA population of a second cell type. For example, one cDNA probe may correspond to a total cell cDNA probe of a cell type or tissue derived from a control subject, while the second cDNA probe may correspond to a total cell cDNA probe of the same cell type or

- 49 -

tissue derived from an experimental subject. Those clones which hybridize to one probe but not to the other potentially represent clones derived from genes differentially expressed in the cell type of interest in control versus experimental subjects.

Subtractive hybridization techniques generally involve the isolation of mRNA taken from two different sources, *e.g.*, control and experimental tissue or cell type, the hybridization of the mRNA or single-stranded cDNA reverse-transcribed from the isolated mRNA, and the removal of all hybridized, and therefore double-stranded, sequences. The remaining non-hybridized, single-stranded cDNAs, potentially represent clones derived from genes that are differentially expressed in the two mRNA sources. Such single-stranded cDNAs are then used as the starting material for the construction of a library comprising clones derived from differentially expressed genes.

The differential display technique describes a procedure, utilizing the well known polymerase chain reaction (PCR; the experimental embodiment set forth in Mullis, U.S. Patent 4,683,202), which allows for the identification of sequences derived from genes which are differentially expressed. First, isolated RNA is reverse-transcribed into single-stranded cDNA, utilizing standard techniques which are well known to those of skill in the art. Primers for the reverse transcriptase reaction may include, but are not limited to, oligo dT-containing primers.

Next, this technique uses pairs of PCR primers, as described below, which allow for the amplification of clones representing a random subset of the RNA transcripts present within any given cell. Utilizing different pairs of primers allows each of the mRNA transcripts present in a cell to be amplified. Among such amplified transcripts may be identified those which have been produced from differentially expressed genes.

- 50 -

The 3' oligonucleotide primer of the primer pairs may contain an oligo dT stretch of 10-13, preferably 11, dT nucleotides at its 5' end, which hybridizes to the poly(A) tail of mRNA or to the complement of a cDNA reverse transcribed from an mRNA poly(A) tail. Second, in order to increase the specificity of the 3' primer, the primer may contain one or more, preferably two, additional nucleotides at its 3' end. Because, statistically, only a subset of the mRNA derived sequences present in the sample of interest will hybridize to such primers, the additional nucleotides allow the primers to amplify only a subset of the mRNA derived sequences present in the sample of interest. This is preferred in that it allows more accurate and complete visualization and characterization of each of the bands representing amplified sequences.

The 5' primer may contain a nucleotide sequence expected, statistically, to have the ability to hybridize to cDNA sequences derived from the tissues of interest. The nucleotide sequence may be an arbitrary one, and the length of the 5' oligonucleotide primer may range from about 9 to about 15 nucleotides, with about 13 nucleotides being preferred. Arbitrary primer sequences cause the lengths of the amplified partial cDNAs produced to be variable, thus allowing different clones to be separated by using standard denaturing sequencing gel electrophoresis.

PCR reaction conditions should be chosen which optimize amplified product yield and specificity, and, additionally, produce amplified products of lengths which may be resolved utilizing standard gel electrophoresis techniques. Such reaction conditions are well known to those of skill in the art, and important reaction parameters include, for example, length and nucleotide sequence of oligonucleotide primers as discussed above, and annealing and elongation step temperatures and reaction times.

The pattern of clones resulting from the reverse transcription and amplification of the mRNA of two different cell types is displayed via sequencing gel electrophoresis and compared. Differentially expressed genes are indicated by differences in the two banding patterns.

- 51 -

Once potentially differentially expressed gene sequences have been identified via bulk techniques such as, for example, those described above, the differential expression of such putatively differentially expressed genes should be corroborated. Corroboration may be accomplished via, for example, such well known techniques as Northern analysis, quantitative RT PCR or RNase protection. Upon corroboration, the differentially expressed genes may be further characterized, and may be identified as target and/or fingerprint genes, as discussed below.

Amplified sequences of differentially expressed genes obtained through, for example, differential display may be used to isolate full length clones of the corresponding gene. The full length coding portion of the gene may readily be isolated, without undue experimentation, by molecular biological techniques well known in the art. For example, the isolated differentially expressed amplified fragment may be labeled and used to screen a cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

PCR technology may also be utilized to isolate full length cDNA sequences. As described above, the isolated, amplified gene fragments obtained through differential display have 5' terminal ends at some random point within the gene and usually have 3' terminal ends at a position corresponding to the 3' end of the transcribed portion of the gene. Once nucleotide sequence information from an amplified fragment is obtained, the remainder of the gene (i.e., the 5' end of the gene, when utilizing differential display) may be obtained using, for example, RT-PCR.

In one embodiment of such a procedure for the identification and cloning of full length gene sequences, RNA may be isolated, following standard procedures, from an appropriate tissue or cellular source. A reverse transcription reaction may then be performed on the RNA using an oligonucleotide primer complimentary to the mRNA that corresponds to the amplified fragment, for the priming of first strand synthesis. Because the primer is anti-parallel to the mRNA, extension will proceed toward the

- 52 -

5' end of the mRNA. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Using the two primers, the 5' portion of the gene is amplified using PCR. Sequences obtained may then be isolated and recombined with previously isolated sequences to generate a full-length cDNA of the differentially expressed genes of the invention. For a review of cloning strategies and recombinant DNA techniques, see e.g., Sambrook *et al.*, 1989, and Ausubel *et al.*, 1989.

Identification of Pathway Genes

Methods are described herein for the identification of pathway genes. "Pathway gene" refers to a gene whose gene product exhibits the ability to interact with gene products involved in cancer, COPD, CNS disorders, or osteoporosis. A pathway gene may be differentially expressed and, therefore, may have the characteristics of a target and/or fingerprint gene.

Any method suitable for detecting protein-protein interactions may be employed for identifying pathway gene products by identifying interactions between gene products and gene products known to be involved in cancer, COPD, CNS disorders, or osteoporosis. Such known gene products may be cellular or extracellular proteins. Those gene products which interact with such known gene products represent pathway gene products and the genes which encode them represent pathway genes.

Among the traditional methods which may be employed are co-immunoprecipitation, crosslinking and co-purification through gradients or chromatographic columns. Utilizing procedures such as these allows for the identification of pathway gene products. Once identified, a pathway gene product may be used, in conjunction with standard techniques, to identify its corresponding pathway gene. For example, at least a portion of the amino acid sequence of the pathway gene product may be ascertained using techniques well known to those of skill in the art, such as via the

- 53 -

Edman degradation technique (see, *e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, W. H. Freeman & Co., N.Y., pp.34-49, 1983). The amino acid sequence obtained may be used as a guide for the generation of oligonucleotide mixtures that can be used to screen for pathway gene sequences. Screening made be accomplished, for example, by standard hybridization or PCR techniques. Techniques for the generation of oligonucleotide mixtures and the screening are well-known. (see, *e.g.*, Ausubel, 1989, and Innis *et al.*, eds., PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS, 1990, Academic Press, Inc., New York).

Methods may be employed which result in the simultaneous identification of pathway genes which encode the protein interacting with a protein involved in cancer, COPD, CNS disorders, or osteoporosis. These methods include, for example, probing expression libraries with labeled protein known or suggested to be involved in such disorders, using this protein in a manner similar to the well known technique of antibody probing of λgt11 libraries.

One method which detects protein interactions *in vivo*, the two-hybrid system, is described in detail for illustration only and not by way of limitation. One version of this system is been described in Chien *et al.*, 1991, *Proc. Natl. Acad. Sci. U.S.A. 88*, 9578-82, 1991, and is commercially available from Clontech (Palo Alto, Calif.). Briefly, utilizing such a system, plasmids are constructed that encode two hybrid proteins: one consists of the DNA-binding domain of a transcription activator protein fused to a known protein, in this case, a protein known to be involved in body weight disorders and or processes relevant to appetite and/or weight regulation, and the other consists of the transcription activator protein's activation domain fused to an unknown protein that is encoded by a cDNA which has been recombined into this plasmid as part of a cDNA library. The plasmids are transformed into a strain of the yeast Saccharomyces cerevisiae that contains a reporter gene (*e.g.*, lacZ) whose regulatory region contains the transcription activator's binding sites. Either hybrid protein alone cannot activate transcription of the reporter gene: the DNA-binding domain hybrid cannot because it does not provide activation function and the

- 54 -

activation domain hybrid cannot because it cannot localize to the activator's binding sites. Interaction of the two hybrid proteins reconstitutes the functional activator protein and results in expression of the reporter gene, which is detected by an assay for the reporter gene product.

The two-hybrid system or related methodology may be used to screen activation domain libraries for proteins that interact with a known "bait" gene product. By way of example, and not by way of limitation, gene products known to be involved in body weight disorders and/or appetite or body weight regulation may be used as the bait gene products. These include but are not limited to the intracellular domain of receptors for such hormones as neuropeptide Y, galanin, interostatin, insulin, and CCK. Total genomic or cDNA sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a hybrid of the bait gene product fused to the DNA-binding domain are cotransformed into a yeast reporter strain, and the resulting transformants are screened for those that express the reporter gene. For example, and not by way of limitation, the bait gene can be cloned into a vector such that it is translationally fused to the DNA encoding the DNA-binding domain of the GAL4 protein. These colonies are purified and the library plasmids responsible for reporter gene expression are isolated. DNA sequencing is then used to identify the proteins encoded by the library plasmids.

A cDNA library of the cell line from which proteins that interact with bait gene product are to be detected can be made using methods routinely practiced in the art. According to the particular system described herein, for example, the cDNA fragments can be inserted into a vector such that they are translationally fused to the activation domain of GAL4. This library can be co-transformed along with the bait gene-GAL4 fusion plasmid into a yeast strain which contains a lacZ gene driven by a promoter which contains GAL4 activation sequence. A cDNA encoded protein, fused to GAL4 activation domain, that interacts with bait gene product will reconstitute an active GAL4 protein and thereby drive expression of the lacZ gene. Colonies which express lacZ can be detected by their blue color in the presence of X-

- 55 -

gal. The cDNA can then be purified from these strains, and used to produce and isolate the bait gene-interacting protein using techniques routinely practiced in the art. Once a pathway gene has been identified and isolated, it may be further characterized, as described below..

Characterization of Differentially Expressed and Pathway Genes

Differentially expressed and pathway genes, such as those identified via the methods discussed above, as well as genes identified by alternative means, may be further characterized by utilizing, for example, methods such as those discussed herein. Such genes will be referred to herein as "identified genes." Analyses such as those described herein, yield information regarding the biological function of the identified genes. An assessment of the biological function of the differentially expressed genes, in addition, will allow for their designation as target and/or fingerprint genes.

Specifically, any of the differentially expressed genes whose further characterization indicates that a modulation of the gene's expression or a modulation of the gene product's activity may ameliorate any of the disorders of interest will be designated "target genes," as defined above. Such target genes and target gene products, along with those discussed below, will constitute the focus of the compound discovery strategies discussed below. Further, such target genes, target gene products and/or modulating compounds can be used as part of the treatment methods described below.

Any of the differentially expressed genes whose further characterization indicates that such modulations may not positively affect body weight disorders of interest, but whose expression pattern contributes to a gene expression "fingerprint" pattern correlative of, for example, a malignant state will be designated a "fingerprint gene." It should be noted that each of the target genes may also function as fingerprint genes, as well as may all or a portion of the pathway genes.

- 56 -

Pathway genes may also be characterized according to techniques such as those described herein. Those pathway genes which yield information indicating that they are differentially expressed and that modulation of the gene's expression or a modulation of the gene product's activity may ameliorate any of the disorders of interest will be also be designated "target genes." Such target genes and target gene products, along with those discussed above, will constitute the focus of the compound discovery strategies discussed below and can be used as part of treatment methods.

Characterization of one or more of the pathway genes may reveal a lack of differential expression, but evidence that modulation of the gene's activity or expression may, nonetheless, ameliorate symptoms. In such cases, these genes and gene products would also be considered a focus of the compound discovery strategies. In instances wherein a pathway gene's characterization indicates that modulation of gene expression or gene product activity may not positively affect disorders of interest, but whose expression is differentially expressed and contributes to a gene expression fingerprint pattern correlative of, for example, cancer, such pathway genes may additionally be designated as fingerprint genes.

A variety of techniques can be utilized to further characterize the identified genes. First, the nucleotide sequence of the identified genes, which may be obtained by utilizing standard techniques well known to those of skill in the art, may, for example, be used to reveal homologies to one or more known sequence motifs which may yield information regarding the biological function of the identified gene product.

Second, an analysis of the tissue and/or cell type distribution of the mRNA produced by the identified genes may be conducted, utilizing standard techniques well known to those of skill in the art. Such techniques may include, for example, Northern, RNase protection and RT-PCR analyses. Such analyses provide information as to, for example, whether the identified genes are expressed in tissues or cell types expected

- 57 -

to contribute to the disorders of interest. Such analyses may also provide quantitative information regarding steady state mRNA regulation, yielding data concerning which of the identified genes exhibits a high level of regulation in, preferably, tissues which may be expected to contribute to the disorders of interest. Additionally, standard *in situ* hybridization techniques may be utilized to provide information regarding which cells within a given tissue express the identified gene. Such an analysis may provide information regarding the biological function of an identified gene relative to a given disorder in instances wherein only a subset of the cells within the tissue is thought to be relevant to the body weight disorder.

Third, the sequences of the identified genes may be used, utilizing standard techniques, to place the genes onto genetic maps, *e.g.*, mouse (Copeland and Jenkins, *Trends in Genetics 7*, 113-18, 1991) and human genetic maps (Cohen *et al.*, *Nature 366*, 698-701, 1993). Such mapping information may yield information regarding the genes' importance to human disease by, for example, identifying genes which map within genetic regions to which known genetic disorders map.

Fourth, the biological function of the identified genes may be more directly assessed by utilizing relevant *in vivo* and *in vitro* systems. *In vivo* systems may include, but are not limited to, animal systems which naturally exhibit symptoms of interest, or ones which have been engineered to exhibit such symptoms. Further, such systems may include systems for the further characterization of CNS disorders, cancer, osteoporosis, and/or COPD and may include, but are not limited to, naturally occurring and transgenic animal systems. *In vitro* systems may include, but are not limited to, cell-based systems comprising cell types known or suspected of contributing to the disorder of interest. Such cells may be wild type cells, or may be non-wild type cells containing modifications known to, or suspected of, contributing to the disorder of interest.

In further characterizing the biological function of the identified genes, the expression of these genes may be modulated within the *in vivo* and/or in vitro systems,

- 58 -

i.e., either overexpressed or underexpressed in, for example, transgenic animals and/or cell lines, and its subsequent effect on the system then assayed. Alternatively, the activity of the product of the identified gene may be modulated by either increasing or decreasing the level of activity in the *in vivo* and/or *in vitro* system of interest, and its subsequent effect then assayed.

The information obtained through such characterizations may suggest relevant methods for the treatment of disorders involving the gene of interest. Further, relevant methods for the treatment of such disorders involving the gene of interest may be suggested by information obtained from such characterizations. For example, treatment may include a modulation of gene expression and/or gene product activity. Characterization procedures such as those described herein may indicate where such modulation should involve an increase or a decrease in the expression or activity of the gene or gene product of interest.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions which can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise a membrane-type serine protease polypeptide, membrane-type serine protease polynucleotide, antibodies which specifically bind to a membrane-type serine protease polypeptide, or mimetics, agonists, antagonists, or inhibitors of a membrane-type serine protease polypeptide. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries

- 59 -

which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as

glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic Indications and Methods

Chronic obstructive pulmonary (or airways) disease (COPD) is a condition defined physiologically as airflow obstruction that generally results from a mixture of emphysema and peripheral airway obstruction due to chronic bronchitis (Senior & Shapiro, *Pulmonary Diseases and Disorders*, 3d ed., New York, McGraw-Hill, 1998, pp. 659-681, 1998; Barnes, *Chest 117*, 10S-14S, 2000). Emphysema is characterized by destruction of alveolar walls leading to abnormal enlargement of the air spaces of the lung. Chronic bronchitis is defined clinically as the presence of chronic productive cough for three months in each of two successive years. In COPD, airflow obstruction is usually progressive and is only partially reversible. By far the most important risk factor for development of COPD is cigarette smoking, although the disease does occur in non-smokers.

Chronic inflammation of the airways is a key pathological feature of COPD (Senior & Shapiro, 1998). The inflammatory cell population comprises increased numbers of macrophages, neutrophils, and $CD8^+$ lymphocytes. Inhaled irritants, such as cigarette smoke, activate macrophages which are resident in the respiratory tract, as well as epithelial cells leading to release of chemokines (*e.g.*, interleukin-8) and other chemotactic factors. These chemotactic factors act to increase the neutrophil/monocyte trafficking from the blood into the lung tissue and airways. Neutrophils and monocytes recruited into the airways can release a variety of potentially damaging mediators such as proteolytic enzymes and reactive oxygen

- 62 -

species. Matrix degradation and emphysema, along with airway wall thickening, surfactant dysfunction, and mucus hypersecretion, all are potential sequelae of this inflammatory response that lead to impaired airflow and gas exchange.

5   COPD is characterized by damage to the lung extracellular matrix and emphysema can be viewed as the pathologic process that affects the lung parenchyma. This process eventually leads to the destruction of the airway walls resulting in permanent airspace enlargement (Senior and Shapiro, in PULMONARY DISEASES AND DISORDERS, 3rd ed., New York, McGraw-Hill, 1998, pp. 659 – 681, 1998). The

10   observation that inherited deficiency of $\alpha 1$-antitrypsin ($\alpha 1$-AT), the primary inhibitor of neutrophil elastase, predisposes individuals to early onset emphysema, and that intrapulmonary instillation of elastolytic enzymes in experimental animals causes emphysema, led to the elastase:antielastase hypothesis for the pathogenesis of emphysema (Eriksson, *Acta Med. Scand. 177(Suppl.)*, 432, 1965, Gross, *J. Occup.*

15   *Med.* 6, 481-84, 1964). This in turn led to the concept that destruction of elastin in the lung parenchyma is the basis of the development of emphysema.

A broad range of immune and inflammatory cells including neutrophils, macrophages, T lymphocytes and eosinophils contain proteolytic enzymes that could

20   contribute to the destruction of lung extracellular matrix (Shapiro, 1999). In addition, a number of different classes of proteases have been identified that have the potential to contribute to lung matrix destruction. These include serine proteases, matrix metalloproteinases and cysteine proteases. Of these classes of enzymes, a number can hydrolyze elastin and have been shown to be elevated in COPD patients

25   (neutrophil elastase, MMP-2, 9, 12) (Culpitt *et al.*, *Am. J. Respir. Crit. Care Med. 160*, 1635-39, 1999, Shapiro, *Am. J. Crit. Care Med. 160 (5)*, S29 – S32,1999).

Membrane-type serine protease is a human ortholog of mouse epithin, a type II membrane serine protease (Kim, et al. Immunogenetics 49, 420-428, 1999).

30   Although the function of epithin is not known, it is likely involved in the processing of peptides or proteins at cell surfaces. Furthermore, membrane-type serine protease

- 63 -

shows high homology to human matriptase, an enzyme that may contribute to the degradation of extracelluar matrix during tumor progression (Lin, et al. J. Biol. Chem. 274, 18231-18236, 1999). The expression of membrane-like serine protease in the lung suggests that it is involved in tissue remodelling, and that dysfunction or dysregulation of the protease plays a significant role in the destruction of the lung matrix in diseases such as COPD. Membrane-type serine protease, therefore, represents a potential therapeutic target for COPD. Regulation of human membrane-type serine protease activity can be used to treat COPD.

Molecules which regulate the expression or activity of human membrane-type serine protease can be used to treat cancer. Cancer is a disease fundamentally caused by oncogenic cellular transformation. There are several hallmarks of transformed cells that distinguish them from their normal counterparts and underlie the pathophysiology of cancer. These include uncontrolled cellular proliferation, un-responsiveness to normal death-inducing signals (immortalization), increased cellular motility and invasiveness, increased ability to recruit blood supply through induction of new blood vessel formation (angiogenesis), genetic instability, and dysregulated gene expression. Various combinations of these aberrant physiologies, along with the acquisition of drug-resistance frequently lead to an intractable disease state in which organ failure and patient death ultimately ensue.

Most standard cancer therapies target cellular proliferation and rely on the differential proliferative capacities between transformed and normal cells for their efficacy. This approach is hindered by the facts that several important normal cell types are also highly proliferative and that cancer cells frequently become resistant to these agents. Thus, the therapeutic indices for traditional anti-cancer therapies rarely exceed 2.0.

The advent of genomics-driven molecular target identification has opened up the possibility of identifying new cancer-specific targets for therapeutic intervention that will provide safer, more effective treatments for cancer patients. Thus, newly

- 64 -

discovered tumor-associated genes and their products can be tested for their role(s) in disease and used as tools to discover and develop innovative therapies. Genes playing important roles in any of the physiological processes outlined above can be characterized as cancer targets.

Genes or gene fragments identified through genomics can readily be expressed in one or more heterologous expression systems to produce functional recombinant proteins. These proteins are characterized *in vitro* for their biochemical properties and then used as tools in high-throughput molecular screening programs to identify chemical modulators of their biochemical activities. Agonists and/or antagonists of target protein activity can be identified in this manner and subsequently tested in cellular and *in vivo* disease models for anti-cancer activity. Optimization of lead compounds with iterative testing in biological models and detailed pharmacokinetic and toxicological analyses form the basis for drug development and subsequent testing in humans.

The human membrane-type serine protease gene provides a therapeutic target for decreasing extracellular matrix degradation, in particular for treating or preventing metastatic cancer. For example, blocking a fibronectin domain of human ephrin-like serine protease can suppress or prevent migration or metastasis of tumor cells in response to fibronectin (9, 10). Cancers whose metastasis can be suppressed according to the invention include adenocarcinoma, melanoma, cancers of the adrenal gland, bladder, bone, breast, cervix, gall bladder, liver, lung, ovary, pancreas, prostate, testis, and uterus. Circulating tumor cells arrested in the capillary beds of different organs must invade the endothelial cell lining and degrade its underlying basement membrane (BM) in order to invade into the extravascular tissue(s) where they establish metastasis (1, 2). Metastatic tumor cells often attach at or near the intercellular junctions between adjacent endothelial cells. Such attachment of the metastatic cells is followed by rupture of the junctions, retraction of the endothelial cell borders and migration through the breach in the endothelium toward the exposed underlying BM (1, 11).

Once located between endothelial cells and the BM, the invading cells must degrade the subendothelial glycoproteins and proteoglycans of the BM in order to migrate out of the vascular compartment. Several cellular enzymes (*e.g.*, collagenase IV, plasminogen activator, cathepsin B, elastase) are thought to be involved in degradation of BM (2, 11). Suppression of human membrane-type serine protease activity therefore can be used to suppress tumor cell invasion and metastasis.

Regulatory molecules of human membrane-type serine protease also can be used to treat osteoporosis. Osteoporosis is a disease characterized by low bone mass and microarchitectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk. It is the most common human metabolic bone disorder. Established osteoporosis includes the presence of fractures. Bone turnover occurs by the action of two major effector cell types within bone: the osteoclast, which is responsible for bone resorption, and the osteoblast, which synthesizes and mineralizes bone matrix. The actions of osteoclasts and osteoblasts are highly co-ordinated. Osteoclast precursors are recruited to the site of turnover; they differentiate and fuse to form mature osteoclasts which then resorb bone. Attached to the bone surface, osteoclasts produce an acidic microenvironment in a tightly defined junction between the specialized osteoclast border membrane and the bone matrix, thus allowing the localized solubilization of bone matrix. This in turn facilitate the proteolysis of demineralized bone collagen. Matrix degradation is thought to release matrix-associated growth factor and cytokines, which recruit osteoblasts in a temporally and spatially controlled fashion. Osteoblasts synthesize and secrete new bone matrix proteins, and subsequently mineralize this new matrix. In the normal skeleton this is a physiological process which does not result in a net change in bone mass. In pathological states, such as osteoporosis, the balance between resorption and formation is altered such that bone loss occurs. See WO 99/45923.

The osteoclast itself is the direct or indirect target of all currently available osteoporosis agents with the possible exception of fluoride. Antiresorptive therapy prevents further bone loss in treated individuals. Osteoblasts are derived from multi-potent stem cells which reside in bone marrow and also gives rise to adipocytes, chondrocytes, fibroblasts and muscle cells. Selective enhancement of osteoblast activity is a highly desirable goal for osteoporosis therapy since it would result in an increase in bone mass, rather than a prevention of further bone loss. An effective anabolic therapy would be expected to lead to a significantly greater reduction in fracture risk than currently available treatments.

The agonists or antagonists to the newly discovered polypeptides may act as antiresorptive by directly altering the osteoclast differentiation, osteoclast adhesion to the bone matrix or osteoclast function of degrading the bone matrix. The agonists or antagonists could indirectly alter the osteoclast function by interfering in the synthesis and/or modification of effector molecules of osteoclast differentiation or function such as cytokines, peptide or steroid hormones, proteases, etc.

The agonists or antagonists to the newly discovered polypeptides may act as anabolics by directly enhancing the osteoblast differentiation and /or its bone matrix forming function. The agonists or antagonists could also indirectly alter the osteoblast function by enhancing the synthesis of growth factors, peptide or steroid hormones or decreasing the synthesis of inhibitory molecules.

The agonists and antagonists may be used to mimic, augment or inhibit the action of the newly discovered polypeptides which may be useful to treat osteoporosis, Paget's disease, degradation of bone implants particularly dental implants.

It is also possible that human membrane-type serine protease activity can be used to degrade proteins involved in neurodegenerative diseases, for example, prion protein amyloid plaques of Genstmann-Straussler Syndrome, Creutzfeldt-Jakob disease, and Scrapie. CNS disorders which may be treated include brain injuries, cerebrovascular

- 67 -

diseases and their consequences, Parkinson's disease, corticobasal degeneration, motor neuron disease, dementia, including ALS, multiple sclerosis, traumatic brain injury, stroke, post-stroke, post-traumatic brain injury, and small-vessel cerebrovascular disease. Dementias, such as Alzheimer's disease, vascular dementia, dementia with Lewy bodies, frontotemporal dementia and Parkinsonism linked to chromosome 17, frontotemporal dementias, including Pick's disease, progressive nuclear palsy, corticobasal degeneration, Huntington's disease, thalamic degeneration, Creutzfeld-Jakob dementia, HIV dementia, schizophrenia with dementia, and Korsakoff's psychosis also can be treated. Similarly, it may be possible to treat cognitive-related disorders, such as mild cognitive impairment, age-associated memory impairment, age-related cognitive decline, vascular cognitive impairment, attention deficit disorders, attention deficit hyperactivity disorders, and memory disturbances in children with learning disabilities, by regulating the activity of human membrane-type serine protease.

Pain that is associated with CNS disorders also can be treated by regulating the activity of human membrane-type serine protease. Pain which can be treated includes that associated with central nervous system disorders, such as multiple sclerosis, spinal cord injury, sciatica, failed back surgery syndrome, traumatic brain injury, epilepsy, Parkinson's disease, post-stroke, and vascular lesions in the brain and spinal cord (e.g., infarct, hemorrhage, vascular malformation). Non-central neuropathic pain includes that associated with post mastectomy pain, reflex sympathetic dystrophy (RSD), trigeminal neuralgiaradioculopathy, post-surgical pain, HIV/AIDS related pain, cancer pain, metabolic neuropathies (e.g., diabetic neuropathy, vasculitic neuropathy secondary to connective tissue disease), paraneoplastic polyneuropathy associated, for example, with carcinoma of lung, or leukemia, or lymphoma, or carcinoma of prostate, colon or stomach, trigeminal neuralgia, cranial neuralgias, and post-herpetic neuralgia. Pain associated with cancer and cancer treatment also can be treated, as can headache pain (for example, migraine with aura, migraine without aura, and other migraine disorders), episodic

- 68 -

and chronic tension-type headache, tension-type like headache, cluster headache, and chronic paroxysmal hemicrania.

The invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a poly-peptide-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

A reagent which affects membrane-type serine protease activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce membrane-type serine protease activity. The reagent preferably binds to an expression product of a human membrane-type serine protease gene. If the expression product is a polypeptide, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells which have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the lipo-some is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung or liver.

- 69 -

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a tumor cell, such as a tumor cell ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods which are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al. Trends in Biotechnol. 11*, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE

- 70 -

TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem.* *263*, 621-24 (1988); Wu *et al.*, *J. Biol. Chem.* *269*, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* *87*, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem.* *266*, 338-42 (1991).

5   If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads,

10   protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Determination of a Therapeutically Effective Dose

15   The determination of a therapeutically effective dose is well within the capability of those skilled in the art.   A therapeutically effective dose refers to that amount of active ingredient which increases or decreases extracellular matrix degradation relative to that which occurs in the absence of the therapeutically effective dose.

20   For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

25

Therapeutic efficacy and toxicity, *e.g.*, ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it

30   can be expressed as the ratio, LD_{50}/ED_{50}.

- 71 -

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors which can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

Effective *in vivo* dosages of an antibody are in the range of about 5 μg to about 50 μg/kg, about 50 μg to about 5 mg/kg, about 100 μg to about 500 μg /kg of patient body weight, and about 200 to about 250 μg /kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg,

about 1 µg to about 2 mg, about 5 µg to about 500 µg, and about 20 µg to about 100 µg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides which express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a membrane-type serine protease polynucleotide or activity of a membrane-type serine protease polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a membrane-type serine protease polynucleotide or the activity of a membrane-type serine protease polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to membrane-type serine protease-specific mRNA, quantitative RT-PCR, immunologic detection of a membrane-type serine protease polypeptide, or measurement of membrane-type serine protease activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

The above disclosure generally describes the present invention, and all patents and patent applications cited in this disclosure are expressly incorporated herein. A more complete understanding can be obtained by reference to the following specific examples which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1

Detection of membrane-type serine protease activity

The polynucleotide of SEQ ID NO: 6 is inserted into the expression vector pCEV4 and the expression vector pCEV4-membrane-type serine protease polypeptide obtained is transfected into human embryonic kidney 293 cells. From these cells extracts are obtained and the membrane-type serine protease activity is measured using thiobenzylester substrates, as described in U.S. Patent 5,500,344. For monitoring enzyme activities from granules and column fractions, assays are performed at room temperature using 0.5 mM 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) (Sigma) to detect the HSBzl leaving group (ϵ_{410} =13600 M^{-1} cm^{-1}).

BLT-esterase activity is estimated using a microtiter assay (Green and Shaw, *Anal. Biochem.* **93**, 223-226, 1979). Briefly, 50 µl of sample is added to 100 µl of 1 mM DTNB, made up in 10 mM HEPES, 1 mM $CaCl_2$, 1 mM $MgCl_2$, pH 7.2. The reaction is initiated by the addition of 50 µl of BLT (Sigma) to give a final concentration of 500 µM. For Metase determinations, 50 µl of dilutions of the sample in 0.1 M HEPES, 0.05 M $CaCl_2$, pH 7.5, are added to 100 µl of 1 mM DTNB, and the reaction is initiated by the addition of 50 µl of Boc-Ala-Ala-Met-S Benzyl (Bzl) to give a final concentration of 150 µM. The duration of the assay depends on color development, the rate of which is measured ($O.D._{410}$) on a Dynatech MR 5000 microplate reader. Controls of sample and DTNB alone or DTNB and substrate alone are run.

For more sensitive comparisons of enzymatic activities, peptide thiobenzyl ester substrates are used to measure protease activities. The chymase substrate Suc-Phe-Leu-Phe-SBzl is purchased from BACHEM Bioscience Inc., Philadelphia, Pa. Z-Arg-SBzl (the tryptase substrate, Kam *et al., J. Biol. Chem.* **262**, 3444-3451, 1987); Boc-Ala-Ala-AA-SBzl (AA=Asp, Met, Leu, Nle, or Ser), and Suc-Ala-Ala-Met-SBzl (Odake *et al, Biochemistry 30*, 2217-2227, 1991); Harper *et al., Biochemistry 23*, 2995-3002, 1984) are synthesized previously.

Boc-Ala-Ala-Asp-SBzl is the substrate for Asp-ase and peptide thiobenzyl esters containing Met, Leu or Nle are substrates for Met-ase SP. Assays are performed at room temperature in 0.1 M, HEPES buffer, pH 7.5, containing 0.01 M $CaCl_2$ and 8% Me_2O using 0.34 mM 4,4'-dithiodipyridine (Aldrithiol-4, Aldrich Chemical Co.,

5   Milwaukee, Wis.) to detect HSBzl leaving group that reacts with 4,4'-dithiodipyridine to release thiopyridone ($\epsilon324$=19800 M^{-1} cm^{-1}, Grasetti and Murray, *Arch. Biochem. Biophys. 119*, 41-49, 1967). The initial rates are measured at 324 nm using a Beckman 35 spectrophotometer when 10-25 µl of an enzyme stock solution is added to a cuvette containing 2.0 ml of buffer, 150 µl of

10   4,4'-dithiodipyridine, and 25 µl of substrate. The same volume of substrate and 4,4'-dithiodipyridine are added to the reference cell in order to compensate for the background hydrolysis rate of the substrates. Initial rates are measured in duplicate for each substrate concentration and are averaged in each case. Substrate concentrations are 100-133 µM. It is shown that the polypeptide of SEQ ID NO: 7

15   has a membrane-type serine protease activity.

EXAMPLE 2

Expression of recombinant membrane-type serine protease

20   The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of a recombinant membrane-type serine protease in yeast. The encoding DNA sequence is derived from the nucleotide sequence shown in SEQ ID NO:6. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5'-end an initiation codon

25   and at its 3'-end an enterokinase cleavage site, a His6 reporter tag, and a termination codon. Moreover, at both termini recognition sequences for restriction endo-nucleases are added.

After digestion of the multiple cloning site of pPICZ B with the corresponding

30   restriction enzymes, the modified DNA sequence is ligated into pPICZB. This

- 76 -

expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is used to transform the yeast.

The yeast is cultivated under usual conditions in 5 liter shake flasks, and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified membrane-type serine protease is obtained.

EXAMPLE 3

Inhibition of extracellular matrix degradation by antisense oligonucleotides

The cell line used for testing is the human colon cancer cell line HCT116. Cells are cultured in RPMI-1640 with 10-15% fetal calf serum at a concentration of 10,000 cells per milliliter in a tissue culture dish on a substrate comprising extracellular matrix components at 37 °C in a 95% air/5% CO_2 atmosphere.

Phosphorothioate oligoribonucleotides are synthesized on an Applied Biosystems Model 380B DNA synthesizer using phosphoroamidite chemistry. The test oligonucleotide is a sequence of 24 bases: 5'-CGG-GTC-GCA-ACA-CTA-CCT-CAT-TTA-3' (complementary to the nucleotides are positions 1-24 of SEQ ID NO:1). As a control, another (random) sequence 5'-TCA-ACT-GAC-TAG-ATG-TAC-ATG-GAC-3' is used. Following assembly and deprotection, oligonucleotides are ethanol-precipitated twice, dried, and suspended in phosphate buffered saline (PBS) at the desired concentration. Purity of these oligonucleotides is tested by capillary gel electrophoresis and ion exchange HPLC. The purified oligonucleotides are added to the culture medium at a concentration of 10 µM once a day for seven days.

- 77 -

The addition of the test oligonucleotide for seven days results in significantly reduced expression of the human membrane-type serine protease as determined by Western blotting. This effect is not observed with the control oligonucleotide. After 3 to 7 days, the integrity of the extracellular matrix substrate is examined using a microscope. The extracellular matrix in the test dishes is largely intact, whereas in the control dishes, degradation of the extracellular matrix substrate has occurred, indicating that inhibition of human membrane-type serine protease has an effect on the ability of cancer cells to degrade extracellular matrix.

EXAMPLE 4

Identification of a test compound which binds to a membrane-type serine protease polypeptide

Purified membrane-type serine protease polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well micro-titer plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Membrane-type serine protease polypeptides comprise the amino acid sequence shown in SEQ ID NOS:2 or 7. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a membrane-type serine protease polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound which increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound was not incubated is identified as a compound which binds to a membrane-type serine protease polypeptide.

- 78 -

EXAMPLE 5

Identification of a test compound which decreases membrane-type serine protease activity

5   Cellular extracts from the human colon cancer cell line HCT116 are contacted with test compounds from a small molecule library and assayed for membrane-type serine protease activity. Control extracts, in the absence of a test compound, also are assayed. Protease activity can be measured using thiobenzylester substrates, as described in U.S. Patent 5,500,344. For monitoring enzyme activities from granules

10   and column fractions, assays are performed at room temperature using 0.5 mM 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) (Sigma) to detect the HSBzl leaving group (ε_{410} =13600 M^{-1} cm^{-1}).

       BLT-esterase activity is estimated using a microtiter assay (Green and Shaw, *Anal.*

15   *Biochem. 93*, 223-226, 1979). Briefly, 50 µl of sample is added to 100 µl of 1 mM DTNB, made up in 10 mM HEPES, 1 mM CaCl$_2$, 1 mM MgCl$_2$, pH 7.2. The reaction is initiated by the addition of 50 µl of BLT (Sigma) to give a final concentration of 500 µM.   For Metase determinations, 50 µl of dilutions of the sample in 0.1 M HEPES, 0.05 M CaCl$_2$, pH 7.5, are added to 100 µl of 1 mM

20   DTNB, and the reaction is initiated by the addition of 50 µl of Boc-Ala-Ala-Met-S Benzyl (Bzl) to give a final concentration of 150 µM. The duration of the assay depends on color development, the rate of which is measured (O.D.$_{410}$) on a Dynatech MR 5000 microplate reader. Controls of sample and DTNB alone or DTNB and substrate alone are run.

25

       For more sensitive comparisons of enzymatic activities, peptide thiobenzyl ester substrates are used to measure protease activities.   The chymase substrate Suc-Phe-Leu-Phe-SBzl is purchased from BACHEM Bioscience Inc., Philadelphia, Pa. Z-Arg-SBzl (the tryptase substrate, Kam *et al., J. Biol. Chem. 262*, 3444-3451,

30   1987); Boc-Ala-Ala-AA-SBzl (AA=Asp, Met, Leu, Nle, or Ser), and Suc-Ala-Ala-Met-SBzl (Odake *et al, Biochemistry 30*, 2217-2227, 1991); Harper *et*

al., Biochemistry 23, 2995-3002, 1984) are synthesized previously. Boc-Ala-Ala-Asp-SBzl is the substrate for Asp-ase and peptide thiobenzyl esters containing Met, Leu or Nle are substrates for Met-ase SP. Assays are performed at room temperature in 0.1 M, HEPES buffer, pH 7.5, containing 0.01 M $CaCl_2$ and 8%

5   Me_2O using 0.34 mM 4,4'-dithiodipyridine (Aldrithiol-4, Aldrich Chemical Co., Milwaukee, Wis.) to detect HSBzl leaving group that reacts with 4,4'-dithiodipyridine to release thiopyridone ($\epsilon 324=19800$ M^{-1} cm^{-1}, Grasetti and Murray, *Arch. Biochem. Biophys. 119*, 41-49, 1967). The initial rates are measured at 324 nm using a Beckman 35 spectrophotometer when 10-25 µl of an enzyme stock

10   solution is added to a cuvette containing 2.0 ml of buffer, 150 µl of 4,4'-dithiodipyridine, and 25 µl of substrate. The same volume of substrate and 4,4'-dithiodipyridine are added to the reference cell in order to compensate for the background hydrolysis rate of the substrates. Initial rates are measured in duplicate for each substrate concentration and are averaged in each case. Substrate

15   concentrations are 100-133 µM.

A test compound which decreases membrane-type serine protease activity of the extract relative to the control extract by at least 20% is identified as a membrane-type serine protease inhibitor.

20

EXAMPLE 6

Identification of a test compound which decreases membrane-type serine protease gene expression

25   A test compound is administered to a culture of the breast tumor cell line MDA-468 and incubated at 37 °C for 10 to 45 minutes. A culture of the same type of cells incubated for the same time without the test compound provides a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al., Biochem. 18*,

30   5294-99, 1979). Northern blots are prepared using 20 to 30 µg total RNA and hybridized with a ^{32}P-labeled membrane-type serine protease-specific probe at 65 ° C

- 80 -

in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NOS:1 or 6. A test compound which decreases the membrane-type serine protease -specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of membrane-type serine protease gene expression.

EXAMPLE 7

Treatment of a breast tumor with a reagent which specifically binds to a membrane-type serine protease gene product

Synthesis of antisense membrane-type serine protease oligonucleotides comprising at least 11 contiguous nucleotides selected from the complement of SEQ ID NOS:1 or 6 is performed on a Pharmacia Gene Assembler series synthesizer using the phosphoramidite procedure (Uhlmann *et al., Chem. Rev. 90*, 534-83, 1990). Following assembly and deprotection, oligonucleotides are ethanol-precipitated twice, dried, and suspended in phosphate-buffered saline (PBS) at the desired concentration. Purity of these oligonucleotides is tested by capillary gel electrophoreses and ion exchange HPLC. Endotoxin levels in the oligonucleotide preparation are determined using the *Limulus* Amebocyte Assay (Bang, *Biol. Bull. (Woods Hole, Mass.) 105*, 361-362, 1953).

An aqueous composition containing the antisense oligonucleotides at a concentration of 0.1-100 µM is injected directly into a breast tumor with a needle. The needle is placed in the tumors and withdrawn while expressing the aqueous composition within the tumor.

The breast tumor is monitored over a period of days or weeks. Additional injections of the antisense oligonucleotides can be given during that time. Metastasis of the breast tumor is suppressed due to decreased membrane-type serine protease activity of the breast tumor cells.

- 81 -

EXAMPLE 8

Tissue-specific expression of membrane-type serine protease

As a first step to establishing a role for membrane-type serine protease in the
pathogenesis of COPD, expression profiling of the gene was done using real-time
quantitative PCR (TaqMan) with RNA samples isolated from a wide range of human
cells and tissues. Total RNA samples were either purchased from commercial
suppliers or purified in-house. Two panels of RNAs were used for profiling: a whole
body organ panel (Table 1) and a respiratory specific panel (Table 2).

Real-time quantitative PCR. This technique is a development of the kinetic analysis
of PCR first described by Higuchi *et al.* (*BioTechnology* *10*, 413-17, 1992;
BioTechnology *11*, 1026-30, 1993). The principle is that at any given cycle within the
exponential phase of PCR, the amount of product is proportional to the initial
number of template copies. PCR amplification is performed in the presence of an
oligonucleotide probe (TaqMan probe) that is complementary to the target sequence
and labeled with a fluorescent reporter dye and a quencher dye. During the extension
phase of PCR, the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA
polymerase, releasing the fluorophore from the effect of the quenching dye (Holland
et al., *Proc. Natl. Acad. Sci. U.S.A. 88*, 7276-80, 1991). Because the fluorescence
emission increases in direct proportion to the amount of the specific amplified
product, the exponential growth phase of PCR product can be detected and used to
determine the initial template concentration (Heid *et al.*, *Genome Res. 6*, 986-94,
1996, and Gibson *et al.*, *Genome Res. 6*, 995-1001, 1996).

RNA extraction and cDNA preparation. Total RNA from each of the 'in-house'
samples listed in Table 2 was isolated using Qiagen's (Crawley, West Sussex, UK)
RNeasy system according to the manufacturer's protocol. The concentration of
purified RNA was determined using RiboGreen RNA quantitation kit (Molecular
Probes Europe, The Netherlands). RNA concentrations of the samples purchased
from commercial suppliers were also determined using RiboGreen. For the

- 82 -

preparation of cDNA, 1µg of total RNA was reverse transcribed using 200U of SUPERSCRIPT™ II RNaseH⁻ Reverse Transcriptase (Life Technologies, Paisley, UK), 10mM dithiothreitol, 0.5mM of each dNTP, and 5µM random hexamers (PE Applied Biosystems, Warrington, Cheshire, UK) in a final volume of 20µl according to the manufacturer's protocol.

TaqMan quantitative analysis. Specific primers and probe were designed according to the recommendations of PE Applied Biosystems and are listed below:

Forward primer:	5'- TTTCCACCTACGGGATCATCA -3'
Reverse primer:	5'- AGGCATCTCTCTTGCCTGACA -3'
Probe:	5'-(FAM)- TGCCTGCACAGAGCATCCGAGAA -3'

where FAM = 6-carboxy-fluorescein.

Quantitative PCR was performed with 10ng of reverse transcribed RNA from each sample. Each determination was done in duplicate.

The assay reaction mix was as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 900nM forward primer; 900nM reverse primer; 200nM probe; 10ng cDNA; and water to 25µl.

Each of the following steps were carried out once: pre PCR, 2 minutes at 50° C, and 10 minutes at 95°C. The following steps were carried out 40 times: denaturation, 15 seconds at 95°C, annealing/extension, 1 minute at 60°C.

Real-time quantitative PCR was done using an ABI Prism 7700 Sequence Detector. The C_T value generated for each reaction was used to determine the initial template concentration (copy number) by interpolation from a universal standard curve. The level of expression of the target gene in each sample was calculated relative to the sample with the lowest expression of the gene.

- 83 -

The relative expression of membrane-type serine protease across various human tissues is shown in Fig. 13. Abundant expression of the gene was detected in lung, prostate, skeletal muscle, testis, and thymus. The gene was expressed at much lower levels or not at all in the other tissues tested. Of particular interest was the expression of membrane-type serine protease in lung and this was investigated further by analysis of the expression of the gene in some of the constituent cell types of the lung. In these samples, there was prominent expression in epithelial cells, with little or no expression in inflammatory cell types (Fig. 14).

Table 1. Human organ RNA panel used for real-time quantitative PCR.
All samples were obtained from Clontech UK Ltd, Basingstoke, UK.

Tissue	Cat. #
Adrenal gland	Human Panel V, K4004-1
Bone marrow	Human Panel II, K4001-1
Brain	Human Panel I, K4000-1
Colon	Human Panel II, K4001-1
Heart	Human Panel III, K4002-1
Kidney	Human Panel I, K4000-1
Liver	Human Panel I, K4000-1
Lung	Human Panel I, K4000-1
Mammary gland	Human Panel III, K4002-1
Pancreas	Human Panel V, K4004-1
Prostate	Human Panel III, K4002-1
Salivary gland	Human Panel V, K4004-1
Skeletal muscle	Human Panel III, K4002-1
Small intestine	Human Panel II, K4001-1
Spleen	Human Panel II, K4001-1
Stomach	Human Panel II, K4001-1

- 84 -

Tissue	Cat. #
Testis	Human Panel III, K4002-1
Thymus	Human Panel II, K4001-1
Thyroid	Human Panel V, K4004-1
Uterus	Human Panel III, K4002-1

Table 2. **Human respiratory specific RNA panel used for real-time quantitative PCR.**

5

Tissue/cell type	Supplier, cat #
Lung (fetal)	Takara (Japan)
Lung	Clontech, Human Panel I, K4000-1
Trachea	Clontech, Human Panel I, K4000-1
Cultured human bronchial epithelial cells	In-house
Cultured airway smooth muscle cells	In-house
Cultured small airway epithelial cells	In-house
Primary cultured alveolar type II cells	In-house
Cultured H441 cells (Clara-like)	In-house
Freshly isolated polymorphonuclear leuko-cytes (neutrophils)	In-house
Freshly isolated monocytes	In-house
Cultured monocytes (macrophage-like)	In-house

EXAMPLE 9

In vivo testing of compounds/target validation

1. Pain:

5   **Acute Pain**

Acute pain is measured on a hot plate mainly in rats. Two variants of hot plate testing are used: In the classical variant animals are put on a hot surface (52 to 56 °C) and the latency time is measured until the animals show nocifensive behavior, such as

10   stepping or foot licking. The other variant is an increasing temperature hot plate where the experimental animals are put on a surface of neutral temperature. Subsequently this surface is slowly but constantly heated until the animals begin to lick a hind paw. The temperature which is reached when hind paw licking begins is a measure for pain threshold.

15

Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Persistent Pain

20

Persistent pain is measured with the formalin or capsaicin test, mainly in rats. A solution of 1 to 5% formalin or 10 to 100 μg capsaicin is injected into one hind paw of the experimental animal. After formalin or capsaicin application the animals show nocifensive reactions like flinching, licking and biting of the affected paw. The

25   number of nocifensive reactions within a time frame of up to 90 minutes is a measure for intensity of pain.

Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o.,

30   i.t., i.c.v., s.c., intradermal, transdermal) prior to formalin or capsaicin administration.

Neuropathic Pain

- 86 -

Neuropathic pain is induced by different variants of unilateral sciatic nerve injury mainly in rats. The operation is performed under anesthesia. The first variant of sciatic nerve injury is produced by placing loosely constrictive ligatures around the common sciatic nerve. The second variant is the tight ligation of about the half of the diameter of the common sciatic nerve. In the next variant, a group of models is used in which tight ligations or transections are made of either the L5 and L6 spinal nerves, or the L% spinal nerve only. The fourth variant involves an axotomy of two of the three terminal branches of the sciatic nerve (tibial and common peroneal nerves) leaving the remaining sural nerve intact whereas the last variant comprises the axotomy of only the tibial branch leaving the sural and common nerves uninjured. Control animals are treated with a sham operation.

Postoperatively, the nerve injured animals develop a chronic mechanical allodynia, cold allodynioa, as well as a thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA; Electronic von Frey System, Somedic Sales AB, Hörby, Sweden). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy), or by means of a cold plate of 5 to 10 °C where the nocifensive reactions of the affected hind paw are counted as a measure of pain intensity. A further test for cold induced pain is the counting of nocifensive reactions, or duration of nocifensive responses after plantar administration of acetone to the affected hind limb. Chronic pain in general is assessed by registering the circadanian rhythms in activity (Surjo and Arndt, Universität zu Köln, Cologne, Germany), and by scoring differences in gait (foot print patterns; FOOTPRINTS program, Klapdor et al., 1997. A low cost method to analyze footprint patterns. J. Neurosci. Methods 75, 49-54).

Compounds are tested against sham operated and vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Inflammatory Pain

Inflammatory pain is induced mainly in rats by injection of 0.75 mg carrageenan or complete Freund's adjuvant into one hind paw. The animals develop an edema with mechanical allodynia as well as thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy, Paw thermal stimulator, G. Ozaki, University of California, USA). For edema measurement two methods are being used. In the first method, the animals are sacrificed and the affected hindpaws sectioned and weighed. The second method comprises differences in paw volume by measuring water displacement in a plethysmometer (Ugo Basile, Comerio, Italy).

Compounds are tested against uninflamed as well as vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.
Diabetic Neuropathic Pain

Rats treated with a single intraperitoneal injection of 50 to 80 mg/kg streptozotocin develop a profound hyperglycemia and mechanical allodynia within 1 to 3 weeks. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA).

Compounds are tested against diabetic and non-diabetic vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

- 88 -

2. Parkinson's disease

6-Hydroxydopamine (6-OH-DA) Lesion

Degeneration of the dopaminergic nigrostriatal and striatopallidal pathways is the central pathological event in Parkinson's disease. This disorder has been mimicked experimentally in rats using single/sequential unilateral stereotaxic injections of 6-OH-DA into the medium forebrain bundle (MFB).

Male Wistar rats (Harlan Winkelmann, Germany), weighing 200±250 g at the beginning of the experiment, are used. The rats are maintained in a temperature- and humidity-controlled environment under a 12 h light/dark cycle with free access to food and water when not in experimental sessions. The following in vivo protocols are approved by the governmental authorities. All efforts are made to minimize animal suffering, to reduce the number of animals used, and to utilize alternatives to in vivo techniques.

Animals are administered pargyline on the day of surgery (Sigma, St. Louis, MO, USA; 50 mg/kg i.p.) in order to inhibit metabolism of 6-OHDA by monoamine oxidase and desmethylimipramine HCl (Sigma; 25 mg/kg i.p.) in order to prevent uptake of 6-OHDA by noradrenergic terminals. Thirty minutes later the rats are anesthetized with sodium pentobarbital (50 mg/kg) and placed in a stereotaxic frame. In order to lesion the DA nigrostriatal pathway 4 µl of 0.01% ascorbic acid-saline containing 8 µg of 6-OHDA HBr (Sigma) are injected into the left medial fore-brain bundle at a rate of 1 µl /min (2.4 mm anterior, 1.49 mm lateral, -2.7 mm ventral to Bregma and the skull surface). The needle is left in place an additional 5 min to allow diffusion to occur.

Stepping Test

Forelimb akinesia is assessed three weeks following lesion placement using a modified stepping test protocol. In brief, the animals are held by the experimenter

with one hand fixing the hindlimbs and slightly raising the hind part above the surface. One paw is touching the table, and is then moved slowly sideways (5 s for 1 m), first in the forehand and then in the backhand direction. The number of adjusting steps is counted for both paws in the backhand and forehand direction of movement. The sequence of testing is right paw forehand and backhand adjusting stepping, followed by left paw forehand and backhand directions. The test is repeated three times on three consecutive days, after an initial training period of three days prior to the first testing. Forehand adjusted stepping reveals no consistent differences between lesioned and healthy control animals. Analysis is therefore restricted to backhand adjusted stepping.

Balance Test

Balance adjustments following postural challenge are also measured during the stepping test sessions. The rats are held in the same position as described in the stepping test and, instead of being moved sideways, tilted by the experimenter towards the side of the paw touching the table. This maneuver results in loss of balance and the ability of the rats to regain balance by forelimb movements is scored on a scale ranging from 0 to 3. Score 0 is given for a normal forelimb placement. When the forelimb movement is delayed but recovery of postural balance detected, score 1 is given. Score 2 represents a clear, yet insufficient, forelimb reaction, as evidenced by muscle contraction, but lack of success in recovering balance, and score 3 is given for no reaction of movement. The test is repeated three times a day on each side for three consecutive days after an initial training period of three days prior to the first testing.

Staircase Test (Paw Reaching)

A modified version of the staircase test is used for evaluation of paw reaching behavior three weeks following primary and secondary lesion placement. Plexiglass test boxes with a central platform and a removable staircase on each side are used.

- 90 -

The apparatus is designed such that only the paw on the same side at each staircase can be used, thus providing a measure of independent forelimb use. For each test the animals are left in the test boxes for 15 min. The double staircase is filled with 7 x 3 chow pellets (Precision food pellets, formula: P, purified rodent diet, size 45 mg;

5   Sandown Scientific) on each side. After each test the number of pellets eaten (successfully retrieved pellets) and the number of pellets taken (touched but dropped) for each paw and the success rate (pellets eaten/pellets taken) are counted separately. After three days of food deprivation (12 g per animal per day) the animals are tested for 11 days. Full analysis is conducted only for the last five days.

10

MPTP treatment

The neurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydro-pyridine (MPTP) causes degeneration of mesencephalic dopaminergic (DAergic) neurons in rodents,

15   non-human primates, and humans and, in so doing, reproduces many of the symptoms of Parkinson's disease. MPTP leads to a marked decrease in the levels of dopamine and its metabolites, and in the number of dopaminergic terminals in the striatum as well as severe loss of the tyrosine hydroxylase (TH)-immunoreactive cell bodies in the substantia nigra, pars compacta.

20

In order to obtain severe and long-lasting lesions, and to reduce mortality, animals receive single injections of MPTP, and are then tested for severity of lesion 7–10 days later. Successive MPTP injections are administered on days 1, 2 and 3. Animals receive application of 4 mg/kg MPTP hydrochloride (Sigma) in saline once daily. All

25   injections are intraperitoneal (i.p.) and the MPTP stock solution is frozen between injections. Animals are decapitated on day 11.

Immunohistology

30   At the completion of behavioral experiments, all animals are anaesthetized with 3 ml thiopental (1 g/40 ml i.p., Tyrol Pharma). The mice are perfused transcardially with

- 91 -

0.01 M PBS (pH 7.4) for 2 min, followed by 4% paraformaldehyde (Merck) in PBS for 15 min. The brains are removed and placed in 4% paraformaldehyde for 24 h at 4 °C. For dehydration they are then transferred to a 20% sucrose (Merck) solution in 0.1 M PBS at 4 °C until they sink. The brains are frozen in methylbutan at -20 °C for 2 min and stored at -70 °C. Using a sledge microtome (mod. 3800-Frigocut, Leica), 25 μm sections are taken from the genu of the corpus callosum (AP 1.7 mm) to the hippocampus (AP 21.8 mm) and from AP 24.16 to AP 26.72. Forty-six sections are cut and stored in assorters in 0.25 M Tris buffer (pH 7.4) for immunohistochemistry.

A series of sections is processed for free-floating tyrosine hydroxylase (TH) immunohistochemistry. Following three rinses in 0.1 M PBS, endogenous peroxidase activity is quenched for 10 min in 0.3% H_2O_2 ±PBS. After rinsing in PBS, sections are preincubated in 10% normal bovine serum (Sigma) for 5 min as blocking agent and transferred to either primary anti-rat TH rabbit antiserum (dilution 1:2000).

Following overnight incubation at room temperature, sections for TH immunoreactivity are rinsed in PBS (2 x10 min) and incubated in biotinylated anti-rabbit immunoglobulin G raised in goat (dilution 1:200) (Vector) for 90 min, rinsed repeatedly and transferred to Vectastain ABC (Vector) solution for 1 h. 3,.3' -Diaminobenzidine tetrahydrochloride (DAB; Sigma) in 0.1 M PBS, supplemented with 0.005% H_2O_2 , serves as chromogen in the subsequent visualization reaction. Sections are mounted on to gelatin-coated slides, left to dry overnight, counter-stained with hematoxylin dehydrated in ascending alcohol concentrations and cleared in butylacetate. Coverslips are mounted on entellan.

Rotarod Test

We use a modification of the procedure described by Rozas and Labandeira-Garcia (1997), with a CR-1 Rotamex system (Columbus Instruments, Columbus, OH) comprising an IBM-compatible personal computer, a CIO-24 data acquisition card, a

- 92 -

control unit, and a four-lane rotarod unit. The rotarod unit consists of a rotating spindle (diameter 7.3 cm) and individual compartments for each mouse. The system software allows preprogramming of session protocols with varying rotational speeds (0–80 rpm). Infrared beams are used to detect when a mouse has fallen onto the base grid beneath the rotarod. The system logs the fall as the end of the experiment for that mouse, and the total time on the rotarod, as well as the time of the fall and all the set-up parameters, are recorded. The system also allows a weak current to be passed through the base grid, to aid training.

3. Dementia
The object recognition task

The object recognition task has been designed to assess the effects of experimental manipulations on the cognitive performance of rodents. A rat is placed in an open field, in which two identical objects are present. The rats inspects both objects during the first trial of the object recognition task. In a second trial, after a retention interval of for example 24 hours, one of the two objects used in the first trial, the 'familiar' object, and a novel object are placed in the open field. The inspection time at each of the objects is registered. The basic measures in the OR task is the time spent by a rat exploring the two object the second trial. Good retention is reflected by higher exploration times towards the novel than the 'familiar' object.

Administration of the putative cognition enhancer prior to the first trial predominantly allows assessment of the effects on acquisition, and eventually on consolidation processes. Administration of the testing compound after the first trial allows to assess the effects on consolidation processes, whereas administration before the second trial allows to measure effects on retrieval processes.

The passive avoidance task

The passive avoidance task assesses memory performance in rats and mice. The inhibitory avoidance apparatus consists of a two-compartment box with a light compartment and a dark compartment. The two compartments are separated by a guillotine door that can be operated by the experimenter. A threshold of 2 cm separates the two compartments when the guillotine door is raised. When the door is open, the illumination in the dark compartment is about 2 lux. The light intensity is about 500 lux at the center of the floor of the light compartment.

Two habituation sessions, one shock session, and a retention session are given, separated by inter-session intervals of 24 hours. In the habituation sessions and the retention session the rat is allowed to explore the apparatus for 300 sec. The rat is placed in the light compartment, facing the wall opposite to the guillotine door. After an accommodation period of 15 sec. the guillotine door is opened so that all parts of the apparatus can be visited freely. Rats normally avoid brightly lit areas and will enter the dark compartment within a few seconds.

In the shock session the guillotine door between the compartments is lowered as soon as the rat has entered the dark compartment with its four paws, and a scrambled 1 mA footshock is administered for 2 sec. The rat is removed from the apparatus and put back into its home cage. The procedure during the retention session is identical to that of the habituation sessions.

The step-through latency, that is the first latency of entering the dark compartment (in sec.) during the retention session is an index of the memory performance of the animal; the longer the latency to enter the dark compartment, the better the retention is. A testing compound in given half an hour before the shock session, together with 1 mg*kg^{-1} scopolamine. Scopolamine impairs the memory performance during the retention session 24 hours later. If the test compound increases the enter latency

- 94 -

compared with the scopolamine-treated controls, is likely to possess cognition enhancing potential.

The Morris water escape task

5

The Morris water escape task measures spatial orientation learning in rodents. It is a test system that has extensively been used to investigate the effects of putative therapeutic on the cognitive functions of rats and mice. The performance of an animal is assessed in a circular water tank with an escape platform that is submerged 10 about 1 cm below the surface of the water. The escape platform is not visible for an animal swimming in the water tank. Abundant extra-maze cues are provided by the furniture in the room, including desks, computer equipment, a second water tank, the presence of the experimenter, and by a radio on a shelf that is playing softly.

15 The animals receive four trials during five daily acquisition sessions. A trial is started by placing an animal into the pool, facing the wall of the tank. Each of four starting positions in the quadrants north, east, south, and west is used once in a series of four trials; their order is randomized. The escape platform is always in the same position. A trial is terminated as soon as the animal had climbs onto the escape platform or 20 when 90 seconds have elapsed, whichever event occurs first. The animal is allowed to stay on the platform for 30 seconds. Then it is taken from the platform and the next trial is started. If an animal did not find the platform within 90 seconds it is put on the platform by the experimenter and is allowed to stay there for 30 seconds. After the fourth trial of the fifth daily session, an additional trial is given as a probe trial: 25 the platform is removed, and the time the animal spends in the four quadrants is measured for 30 or 60 seconds. In the probe trial, all animals start from the same start position, opposite to the quadrant where the escape platform had been positioned during acquisition.

30 Four different measures are taken to evaluate the performance of an animal during acquisition training: escape latency, traveled distance, distance to platform, and

- 95 -

swimming speed. The following measures are evaluated for the probe trial: time (s) in quadrants and traveled distance (cm) in the four quadrants. The probe trial provides additional information about how well an animal learned the position of the escape platform. If an animal spends more time and swims a longer distance in the quadrant where the platform had been positioned during the acquisition sessions than in any other quadrant, one concludes that the platform position has been learned well.

In order to assess the effects of putative cognition enhancing compounds, rats or mice with specific brain lesions which impair cognitive functions, or animals treated with compounds such as scopolamine or MK-801, which interfere with normal learning, or aged animals which suffer from cognitive deficits, are used.

The T-maze spontaneous alternation task

The T-maze spontaneous alternation task (TeMCAT) assesses the spatial memory performance in mice. The start arm and the two goal arms of the T-maze are provided with guillotine doors which can be operated manually by the experimenter. A mouse is put into the start arm at the beginning of training. The guillotine door is closed. In the first trial, the 'forced trial', either the left or right goal arm is blocked by lowering the guillotine door. After the mouse has been released from the start arm, it will negotiate the maze, eventually enter the open goal arm, and return to the start position, where it will be confined for 5 seconds, by lowering the guillotine door. Then, the animal can choose freely between the left and right goal arm (all guillotine-doors opened) during 14 'free choice' trials. As soon a the mouse has entered one goal arm, the other one is closed. The mouse eventually returns to the start arm and is free to visit whichever go alarm it wants after having been confined to the start arm for 5 seconds. After completion of 14 free choice trials in one session, the animal is removed from the maze. During training, the animal is never handled.

- 96 -

The per-cent alternations out of 14 trials is calculated. This percentage and the total time needed to complete the first forced trial and the subsequent 14 free choice trials (in s) is analyzed. Cognitive deficits are usually induced by an injection of scopolamine, 30 min before the start of the training session. Scopolamine reduced the
5   per-cent alternations to chance level, or below. A cognition enhancer, which is always administered before the training session, will at least partially, antagonize the scopolamine-induced reduction in the spontaneous alternation rate.

EXAMPLE 10

10   *Proliferation inhibition assay: Antisense oligonucleotides suppress the growth of cancer cell lines*

The cell line used for testing is the human colon cancer cell line HCT116. Cells are cultured in RPMI-1640 with 10-15% fetal calf serum at a concentration of 10,000
15   cells per milliliter in a volume of 0.5 ml and kept at 37 °C in a 95% air/5%CO_2 atmosphere.

Phosphorothioate oligoribonucleotides are synthesized on an Applied Biosystems Model 380B DNA synthesizer using phosphoroamidite chemistry. A sequence of 24
20   bases complementary to the nucleotides at position 1 to 24 of SEQ ID NO:1 is used as the test oligonucleotide. As a control, another (random) sequence is used: 5'-TCA ACT GAC TAG ATG TAC ATG GAC-3'. Following assembly and deprotection, oligonucleotides are ethanol-precipitated twice, dried, and suspended in phosphate buffered saline at the desired concentration. Purity of the oligonucleotides is tested
25   by capillary gel electrophoresis and ion exchange HPLC. The purified oligonucleotides are added to the culture medium at a concentration of 10 µM once per day for seven days.

The addition of the test oligonucleotide for seven days results in significantly
30   reduced expression of human membrane-type serine protease as determined by Western blotting. This effect is not observed with the control oligonucleotide. After

- 97 -

3 to 7 days, the number of cells in the cultures is counted using an automatic cell counter. The number of cells in cultures treated with the test oligonucleotide (expressed as 100%) is compared with the number of cells in cultures treated with the control oligonucleotide. The number of cells in cultures treated with the test

5   oligonucleotide is not more than 30% of control, indicating that the inhibition of human membrane-type serine protease has an anti-proliferative effect on cancer cells.

EXAMPLE 11

In vivo testing of compounds/target validation

10   **1.   Acute Mechanistic Assays**

1.1.   Reduction in Mitogenic Plasma Hormone Levels

This non-tumor assay measures the ability of a compound to reduce either the endogenous level of a circulating hormone or the level of hormone produced in

15   response to a biologic stimulus. Rodents are administered test compound (p.o., i.p., i.v., i.m., or s.c.). At a predetermined time after administration of test compound, blood plasma is collected. Plasma is assayed for levels of the hormone of interest. If the normal circulating levels of the hormone are too low and/or variable to provide consistent results, the level of the hormone may be elevated by a pre-treatment with a

20   biologic stimulus (i.e., LHRH may be injected i.m. into mice at a dosage of 30 ng/mouse to induce a burst of testosterone synthesis). The timing of plasma collection would be adjusted to coincide with the peak of the induced hormone response. Compound effects are compared to a vehicle-treated control group. An F-test is preformed to determine if the variance is equal or unequal followed by a

25   Student's t-test. Significance is p value ≤ 0.05 compared to the vehicle control group.

1.2.   Hollow Fiber Mechanism of Action Assay

30   Hollow fibers are prepared with desired cell line(s) and implanted intraperitoneally and/or subcutaneously in rodents. Compounds are administered p.o., i.p., i.v., i.m., or

- 98 -

s.c. Fibers are harvested in accordance with specific readout assay protocol, these may include assays for gene expression (bDNA, PCR, or Taqman), or a specific biochemical activity (i.e., cAMP levels. Results are analyzed by Student's t-test or Rank Sum test after the variance between groups is compared by an F-test, with significance at $p \leq 0.05$ as compared to the vehicle control group.

2. Subacute Functional *In Vivo* Assays

2.1. *Reduction in Mass of Hormone Dependent Tissues*

This is another non-tumor assay that measures the ability of a compound to reduce the mass of a hormone dependent tissue (i.e., seminal vesicles in males and uteri in females). Rodents are administered test compound (p.o., i.p., i.v., i.m., or s.c.) according to a predetermined schedule and for a predetermined duration (i.e., 1 week). At termination of the study, animals are weighed, the target organ is excised, any fluid is expressed, and the weight of the organ is recorded. Blood plasma may also be collected. Plasma may be assayed for levels of a hormone of interest or for levels of test agent. Organ weights may be directly compared or they may be normalized for the body weight of the animal. Compound effects are compared to a vehicle-treated control group. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test. Significance is p value ≤ 0.05 compared to the vehicle control group.

2.2. *Hollow Fiber Proliferation Assay*

Hollow fibers are prepared with desired cell line(s) and implanted intraperitoneally and/or subcutaneously in rodents. Compounds are administered p.o., i.p., i.v., i.m., or s.c. Fibers are harvested in accordance with specific readout assay protocol. Cell proliferation is determined by measuring a marker of cell number (i.e., MTT or LDH). The cell number and change in cell number from the starting inoculum are analyzed by Student's t-test or Rank Sum test after the variance between groups is

compared by an F-test, with significance at $p \leq 0.05$ as compared to the vehicle control group.

2.3.   Anti-angiogenesis Models

2.3.1. Corneal Angiogenesis

Hydron pellets with or without growth factors or cells are implanted into a micropocket surgically created in the rodent cornea. Compound administration may be systemic or local (compound mixed with growth factors in the hydron pellet). Corneas are harvested at 7 days post implantation immediately following intracardiac infusion of colloidal carbon and are fixed in 10% formalin. Readout is qualitative scoring and/or image analysis. Qualitative scores are compared by Rank Sum test. Image analysis data is evaluated by measuring the area of neovascularization (in pixels) and group averages are compared by Student's t-test (2 tail). Significance is $p \leq 0.05$ as compared to the growth factor or cells only group.

2.3.2. Matrigel Angiogenesis

Matrigel, containing cells or growth factors, is injected subcutaneously. Compounds are administered p.o., i.p., i.v., i.m., or s.c. Matrigel plugs are harvested at predetermined time point(s) and prepared for readout. Readout is an ELISA-based assay for hemoglobin concentration and/or histological examination (i.e. vessel count, special staining for endothelial surface markers: CD31, factor-8). Readouts are analyzed by Student's t-test, after the variance between groups is compared by an F-test, with significance determined at $p \leq 0.05$ as compared to the vehicle control group.

- 100 -

3.   **Primary Antitumor Efficacy**

3.1.   Early Therapy Models

3.1.1.   Subcutaneous Tumor

5   Tumor cells or fragments are implanted subcutaneously on Day 0. Vehicle and/or compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule starting at a time, usually on Day 1, prior to the ability to measure the tumor burden. Body weights and tumor measurements are recorded 2-3 times weekly. Mean net body and tumor weights are calculated for each data collection day. Anti-

10   tumor efficacy may be initially determined by comparing the size of treated (T) and control (C) tumors on a given day by a Student's t-test, after the variance between groups is compared by an F-test, with significance determined at $p \leq 0.05$. The experiment may also be continued past the end of dosing in which case tumor measurements would continue to be recorded to monitor tumor growth delay. Tumor

15   growth delays are expressed as the difference in the median time for the treated and control groups to attain a predetermined size divided by the median time for the control group to attain that size. Growth delays are compared by generating Kaplan-Meier curves from the times for individual tumors to attain the evaluation size. Significance is $p \leq 0.05$.

20

3.1.2.   Intraperitoneal/Intracranial Tumor Models

Tumor cells are injected intraperitoneally or intracranially on Day 0. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule

25   starting on Day 1. Observations of morbidity and/or mortality are recorded twice daily. Body weights are measured and recorded twice weekly. Morbidity/mortality data is expressed in terms of the median time of survival and the number of long-term survivors is indicated separately. Survival times are used to generate Kaplan-Meier curves. Significance is $p \leq 0.05$ by a log-rank test compared to the control

30   group in the experiment.

3.2. Established Disease Model

Tumor cells or fragments are implanted subcutaneously and grown to the desired size for treatment to begin. Once at the predetermined size range, mice are randomized into treatment groups. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Tumor and body weights are measured and recorded 2-3 times weekly. Mean tumor weights of all groups over days post inoculation are graphed for comparison. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control groups at the end of treatment. Significance is $p \leq 0.05$ as compared to the control group. Tumor measurements may be recorded after dosing has stopped to monitor tumor growth delay. Tumor growth delays are expressed as the difference in the median time for the treated and control groups to attain a predetermined size divided by the median time for the control group to attain that size. Growth delays are compared by generating Kaplan-Meier curves from the times for individual tumors to attain the evaluation size. Significance is p value $\leq$ 0.05 compared to the vehicle control group.

3.3. Orthotopic Disease Models

3.3.1. Mammary Fat Pad Assay

Tumor cells or fragments, of mammary adenocarcinoma origin, are implanted directly into a surgically exposed and reflected mammary fat pad in rodents. The fat pad is placed back in its original position and the surgical site is closed. Hormones may also be administered to the rodents to support the growth of the tumors. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Tumor and body weights are measured and recorded 2-3 times weekly. Mean tumor weights of all groups over days post inoculation are graphed for comparison. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control

- 102 -

groups at the end of treatment. Significance is $p \leq 0.05$ as compared to the control group.

Tumor measurements may be recorded after dosing has stopped to monitor tumor growth delay. Tumor growth delays are expressed as the difference in the median time for the treated and control groups to attain a predetermined size divided by the median time for the control group to attain that size. Growth delays are compared by generating Kaplan-Meier curves from the times for individual tumors to attain the evaluation size. Significance is p value≤ 0.05 compared to the vehicle control group. In addition, this model provides an opportunity to increase the rate of spontaneous metastasis of this type of tumor. Metastasis can be assessed at termination of the study by counting the number of visible foci per target organ, or measuring the target organ weight. The means of these endpoints are compared by Student's t-test after conducting an F-test, with significance determined at $p \leq 0.05$ compared to the control group in the experiment.

3.3.2. Intraprostatic Assay

Tumor cells or fragments, of prostatic adenocarcinoma origin, are implanted directly into a surgically exposed dorsal lobe of the prostate in rodents. The prostate is externalized through an abdominal incision so that the tumor can be implanted specifically in the dorsal lobe while verifying that the implant does not enter the seminal vesicles. The successfully inoculated prostate is replaced in the abdomen and the incisions through the abdomen and skin are closed. Hormones may also be administered to the rodents to support the growth of the tumors. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Body weights are measured and recorded 2-3 times weekly. At a predetermined time, the experiment is terminated and the animal is dissected. The size of the primary tumor is measured in three dimensions using either a caliper or an ocular micrometer attached to a dissecting scope. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated

and control groups at the end of treatment. Significance is $p \leq 0.05$ as compared to the control group. This model provides an opportunity to increase the rate of spontaneous metastasis of this type of tumor. Metastasis can be assessed at termination of the study by counting the number of visible foci per target organ (i.e.,

5   the lungs), or measuring the target organ weight (i.e., the regional lymph nodes). The means of these endpoints are compared by Student's t-test after conducting an F-test, with significance determined at $p \leq 0.05$ compared to the control group in the experiment.

10   *3.3.3. Intrabronchial Assay*

Tumor cells of pulmonary origin may be implanted intrabronchially by making an incision through the skin and exposing the trachea. The trachea is pierced with the beveled end of a 25 gauge needle and the tumor cells are inoculated into the main

15   bronchus using a flat-ended 27 gauge needle with a 90° bend. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Body weights are measured and recorded 2-3 times weekly. At a predetermined time, the experiment is terminated and the animal is dissected. The size of the primary tumor is measured in three dimensions using either a caliper or an ocular micrometer

20   attached to a dissecting scope. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control groups at the end of treatment. Significance is $p \leq 0.05$ as compared to the control group. This model provides an opportunity to increase the rate of spontaneous metastasis of this type of tumor. Metastasis can be assessed at

25   termination of the study by counting the number of visible foci per target organ (i.e., the contralateral lung), or measuring the target organ weight. The means of these endpoints are compared by Student's t-test after conducting an F-test, with significance determined at $p \leq 0.05$ compared to the control group in the experiment.

3.3.4. Intracecal Assay

Tumor cells of gastrointestinal origin may be implanted intracecally by making an abdominal incision through the skin and externalizing the intestine. Tumor cells are inoculated into the cecal wall without penetrating the lumen of the intestine using a 27 or 30 gauge needle. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Body weights are measured and recorded 2-3 times weekly. At a predetermined time, the experiment is terminated and the animal is dissected. The size of the primary tumor is measured in three dimensions using either a caliper or an ocular micrometer attached to a dissecting scope. An F-test is preformed to determine if the variance is equal or unequal followed by a Student's t-test to compare tumor sizes in the treated and control groups at the end of treatment. Significance is $p \leq 0.05$ as compared to the control group. This model provides an opportunity to increase the rate of spontaneous metastasis of this type of tumor. Metastasis can be assessed at termination of the study by counting the number of visible foci per target organ (i.e., the liver), or measuring the target organ weight. The means of these endpoints are compared by Student's t-test after conducting an F-test, with significance determined at $p \leq 0.05$ compared to the control group in the experiment.

4.   Secondary (Metastatic) Antitumor Efficacy

4.1.   Spontaneous Metastasis

Tumor cells are inoculated s.c. and the tumors allowed to grow to a predetermined range for spontaneous metastasis studies to the lung or liver. These primary tumors are then excised. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule which may include the period leading up to the excision of the primary tumor to evaluate therapies directed at inhibiting the early stages of tumor metastasis. Observations of morbidity and/or mortality are recorded daily. Body weights are measured and recorded twice weekly. Potential endpoints include

- 105 -

survival time, numbers of visible foci per target organ, or target organ weight. When survival time is used as the endpoint the other values are not determined. Survival data is used to generate Kaplan-Meier curves. Significance is $p \leq 0.05$ by a log-rank test compared to the control group in the experiment. The mean number of visible tumor foci, as determined under a dissecting microscope, and the mean target organ weights are compared by Student's t-test after conducting an F-test, with significance determined at $p \leq 0.05$ compared to the control group in the experiment for both of these endpoints.

4.2. Forced Metastasis

Tumor cells are injected into the tail vein, portal vein, or the left ventricle of the heart in experimental (forced) lung, liver, and bone metastasis studies, respectively. Compounds are administered p.o., i.p., i.v., i.m., or s.c. according to a predetermined schedule. Observations of morbidity and/or mortality are recorded daily. Body weights are measured and recorded twice weekly. Potential endpoints include survival time, numbers of visible foci per target organ, or target organ weight. When survival time is used as the endpoint the other values are not determined. Survival data is used to generate Kaplan-Meier curves. Significance is $p \leq 0.05$ by a log-rank test compared to the control group in the experiment. The mean number of visible tumor foci, as determined under a dissecting microscope, and the mean target organ weights are compared by Student's t-test after conducting an F-test, with significance at $p \leq 0.05$ compared to the vehicle control group in the experiment for both endpoints.

EXAMPLE 12

Identification of test compound efficacy in a COPD animal model

Guinea pigs are exposed on a single occasion to tobacco smoke for 50 minutes. Animals are sacrificed between 10 minutes and 24 hour following the end of the

- 106 -

exposure and their lungs placed in RNAlater™. The lung tissue is homogenised ,and total RNA was extracted using a Qiagens RNeasy™ Maxi kit. Molecular Probes RiboGreen™ RNA quantitation method is used to quantify the amount of RNA in each sample.

Total RNA is reverse transcribed, and the resultant cDNA is used in a real-time polymerase chain reaction (PCR). The cDNA is added to a solution containing the sense and anti-sense primers and the 6-carboxy-tetramethyl-rhodamine labelled probe of the membrane-type serine protease gene. Cyclophilin is used as the housekeeping gene. The expression of the membrane-type serine protease gene is measured using the TaqMan real-time PCR system that generates an amplification curve for each sample. From this curve a threshold cycle value is calculated: the fractional cycle number at which the amount of amplified target reaches a fixed threshold. A sample containing many copies of the membrane-type serine protease gene will reach this threshold earlier than a sample containing fewer copies. The threshold is set at 0.2, and the threshold cycle C_T is calculated from the amplification curve. The C_T value for the membrane-type serine protease gene is normalised using the C_T value for the housekeeping gene.

Expression of the membrane-type serine protease gene is increased by at least 3-fold between 10 minutes and 3 hours post tobacco smoke exposure compared to air exposed control animals.

Test compounds are evaluated as follows. Animals are pre-treated with a test compound between 5 minutes and 1 hour prior to the tobacco smoke exposure and they are then sacrificed up to 3 hours after the tobacco smoke exposure has been completed. Control animals are pre-treated with the vehicle of the test compound via the route of administration chosen for the test compound. A test compound that reduces the tobacco smoke induced upregulation of membrane-type serine protease gene relative to the expression seen in vehicle treated tobacco smoke exposed

animals is identified as an inhibitor of membrane-type serine protease gene expression.

REFERENCES

1. Nicolson (1988) Organ specificity of tumor metastasis: Role of preferential adhesion, invasion and growth of malignant cells at specific secondary sites. *Cancer Met. Rev. 7*, 143-188.

2. Liotta *et al.* (1983) Tumor invasion and the extracellular matrix. *Lab. Invest. 49*, 639-649.

3. Vlodavsky *et al.* (1987) Endothelial cell-derived basic fibroblast growth factor: Synthesis and deposition into subendothelial extracellular matrix. *Proc. Natl. Acad. Sci. USA 84*, 2292-2296.

4. Folkman *et al.* (1980) A heparin-binding angiogenic protein--basic fibroblast growth factor--is stored within basement membrane. *Am. J. Pathol. 130, 393400.*

5. Cardon-Cardo *et al.* (1990) Expression of basic fibroblast growth factor in normal human tissues. *Lab. Invest. 63*, 832-840.

6. Vlodavsky *et al.* (1991) Extracellular sequestration and release of fibroblast growth factor: a regulatory mechanism? *Trends Biochem. Sci. 16*, 268-271.

7. Vlodavsky *et al.* (1993) Extracellular matrix-bound growth factors, enzymes and plasma proteins. In BASEMENT MEMBRANES: CELLULAR AND MOLECULAR ASPECTS Rohrbach & Timpl, eds., pp327-343. Academic Press Inc., Orlando, Fla.

8. Ross (1993) The pathogenesis of atherosclerosis: a perspective for the 1990s. *Nature (Lond.). 362*, 801-809.

9. McCarthy *et al.* (1986). Human fibronectin contains distinct adhesion- and motility-promoting domains for metastatic melanoma cells. *J. Cell Biol. 102*, 179-88.

- 108 -

10.   Van Muijen *et al.*   (1995) Properties of metastasizing and non-metastasizing human melanoma cells.   *Recent Results in Cancer Research 139,* 104-22.

11.   Price *et al.*   (1997) The Biochemistry of Cancer Dissemination, in *Critical Reviews in Biochemistry and Mol. Biol. 32,* 175-253.

CLAIMS

1.   An isolated polynucleotide encoding a membrane-type serine protease polypeptide and being selected from the group consisting of:

   a)   a polynucleotide encoding a membrane-type serine protease polypeptide comprising an amino acid sequence selected form the group consisting of:

   amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2;

   the amino acid sequence shown in SEQ ID NO: 2;

   amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7; and

   the amino acid sequence shown in SEQ ID NO: 7.

   b)   a polynucleotide comprising the sequence of SEQ ID NO: 1 or 6;

   c)   a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

   d)   a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

   e)   a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a to (d).

2.   An expression vector containing any polynucleotide of claim 1.

3.   A host cell containing the expression vector of claim 2.

4.   A substantially purified membrane-type serine protease polypeptide encoded by a polynucleotide of claim 1.

5.   A method for producing a membrane-type serine protease polypeptide, wherein the method comprises the following steps:

- 110 -

a)   culturing the host cell of claim 3 under conditions suitable for the expression of the membrane-type serine protease polypeptide; and

b)   recovering the membrane-type serine protease polypeptide from the host cell culture.

5

6.   A method for detection of a polynucleotide encoding a membrane-type serine protease polypeptide in a biological sample comprising the following steps:

a)   hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and

10   b)   detecting said hybridization complex.

7.   The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

15   8.   A method for the detection of a polynucleotide of claim 1 or a membrane-type serine protease polypeptide of claim 4 comprising the steps of:

contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the membrane-type serine protease polypeptide.

20   9.   A diagnostic kit for conducting the method of any one of claims 6 to 8.

10.   A method of screening for agents which decrease the activity of a membrane-type serine protease, comprising the steps of:

contacting a test compound with any membrane-type serine protease polypeptide encoded by any polynucleotide of claim 1;

25

detecting binding of the test compound to the membrane-type serine protease polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a membrane-type serine protease.

30

- 111 -

11.   A method of screening for agents which regulate the activity of a membrane-type serine protease, comprising the steps of:

contacting a test compound with a membrane-type serine protease polypeptide encoded by any polynucleotide of claim 1; and

detecting a membrane-type serine protease activity of the polypeptide, wherein a test compound which increases the membrane-type serine protease activity is identified as a potential therapeutic agent for increasing the activity of the membrane-type serine protease, and wherein a test compound which decreases the membrane-type serine protease activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the membrane-type serine protease.

12.   A method of screening for agents which decrease the activity of a membrane-type serine protease, comprising the steps of:

contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of membrane-type serine protease.

13.   A method of reducing the activity of membrane-type serine protease, comprising the steps of:

contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any membrane-type serine protease polypeptide of claim 4, whereby the activity of membrane-type serine protease is reduced.

14.   A reagent that modulates the activity of a membrane-type serine protease polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.

15.   A pharmaceutical composition, comprising:

- 112 -

the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

16.   Use of the pharmaceutical composition of claim 15 for modulating the activity of a membrane-type serine protease in a disease.

17.   Use of claim 16 wherein the disease is cancer, a neurodegenerative disease, osteoporosis or chronic obstructive pulmonary disease.

18.   A cDNA encoding a polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7.

19.   The cDNA of claim 18 which comprises SEQ ID NO: 1 or 6.

20.   The cDNA of claim 18 which consists of SEQ ID NO: 1 or 6.

21.   An expression vector comprising a polynucleotide which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7.

22.   The expression vector of claim 21 wherein the polynucleotide consists of SEQ ID NO: 1 or 6.

23.   A host cell comprising an expression vector which encodes a polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7.

24.   The host cell of claim 23 wherein the polynucleotide consists of SEQ ID NO: 1 or 6.

25.   A purified polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7.

26.   The purified polypeptide of claim 25 which consists of the amino acid sequence shown in SEQ ID NOS: 2 or 7.

27.   A fusion protein comprising a polypeptide having the amino acid sequence shown in SEQ ID NOS: 2 or 7.

28.   A method of producing a polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7, comprising the steps of:
culturing a host cell comprising an expression vector which encodes the polypeptide under conditions whereby the polypeptide is expressed; and isolating the polypeptide.

29.   The method of claim 28 wherein the expression vector comprises SEQ ID NO: 1 or 6.

30.   A method of detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7, comprising the steps of:
hybridizing a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO: 1 or 6 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and
detecting the hybridization complex.

31.   The method of claim 30 further comprising the step of amplifying the nucleic acid material before the step of hybridizing.

32.   A kit for detecting a coding sequence for a polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7, comprising:
a polynucleotide comprising 11 contiguous nucleotides of SEQ ID NO: 1 or 6; and
instructions for the method of claim 30.

33.   A method of detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7, comprising the steps of:
        contacting a biological sample with a reagent that specifically binds to the polypeptide to form a reagent-polypeptide complex; and
        detecting the reagent-polypeptide complex.

34.   The method of claim 33 wherein the reagent is an antibody.

35.   A kit for detecting a polypeptide comprising the amino acid sequence shown in SEQ ID NOS: 2 or 7, comprising:
        an antibody which specifically binds to the polypeptide; and
        instructions for the method of claim 33.

36.   A method of screening for agents which can modulate the activity of a human membrane-type serine protease, comprising the steps of:
        contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NOS: 2 or 7 and (2) the amino acid sequence shown in SEQ ID NOS: 2 or 7; and
        detecting binding of the test compound to the polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential agent for regulating activity of the human membrane-type serine protease.

37.   The method of claim 36 wherein the step of contacting is in a cell.

38.   The method of claim 36 wherein the cell is *in vitro*.

39.   The method of claim 36 wherein the step of contacting is in a cell-free system.

- 115 -

40.   The method of claim 36 wherein the polypeptide comprises a detectable label.

41.   The method of claim 36 wherein the test compound comprises a detectable label.

5

42.   The method of claim 36 wherein the test compound displaces a labeled ligand which is bound to the polypeptide.

43.   The method of claim 36 wherein the polypeptide is bound to a solid support.

10

44.   The method of claim 36 wherein the test compound is bound to a solid support.

45.   A method of screening for agents which modulate an activity of a human membrane-type serine protease, comprising the steps of:

15   contacting a test compound with a polypeptide comprising an amino acid sequence selected from the group consisting of: (1) amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NOS: 2 or 7 and (2) the amino acid sequence shown in SEQ ID NOS:

20   2 or 7; and

detecting an activity of the polypeptide, wherein a test compound which increases the activity of the polypeptide is identified as a potential agent for increasing the activity of the human membrane-type serine protease, and wherein a test compound which decreases the activity of the polypeptide is

25   identified as a potential agent for decreasing the activity of the human membrane-type serine protease.

46.   The method of claim 45 wherein the step of contacting is in a cell.

30   47.   The method of claim 45 wherein the cell is *in vitro*.

48.   The method of claim 45 wherein the step of contacting is in a cell-free system.

49.   A method of screening for agents which modulate an activity of a human membrane-type serine protease, comprising the steps of:

contacting a test compound with a product encoded by a polynucleotide which comprises the nucleotide sequence shown in SEQ ID NO: 1 or 6; and

detecting binding of the test compound to the product, wherein a test compound which binds to the product is identified as a potential agent for regulating the activity of the human membrane-type serine protease.

50.   The method of claim 49 wherein the product is a polypeptide.

51.   The method of claim 49 wherein the product is RNA.

52.   A method of reducing activity of a human membrane-type serine protease, comprising the step of:

contacting a cell with a reagent which specifically binds to a product encoded by a polynucleotide comprising the nucleotide sequence shown in SEQ ID NO: 1 or 6, whereby the activity of a human membrane-type serine protease is reduced.

53.   The method of claim 52 wherein the product is a polypeptide.

54.   The method of claim 53 wherein the reagent is an antibody.

55.   The method of claim 52 wherein the product is RNA.

56.   The method of claim 55 wherein the reagent is an antisense oligonucleotide.

57.   The method of claim 56 wherein the reagent is a ribozyme.

58.   The method of claim 52 wherein the cell is *in vitro*.

59.   The method of claim 52 wherein the cell is *in vivo*.

5   60.   A pharmaceutical composition, comprising:
        a reagent which specifically binds to a polypeptide comprising the amino acid
        sequence shown in SEQ ID NOS: 2 or 7; and
        a pharmaceutically acceptable carrier.

10   61.   The pharmaceutical composition of claim 60 wherein the reagent is an
        antibody.

62.   A pharmaceutical composition, comprising:
        a reagent which specifically binds to a product of a polynucleotide comprising
15   the nucleotide sequence shown in SEQ ID NO: 1 or 6; and
        a pharmaceutically acceptable carrier.

63.   The pharmaceutical composition of claim 62 wherein the reagent is a
        ribozyme.

20
64.   The pharmaceutical composition of claim 62 wherein the reagent is an
        antisense oligonucleotide.

65.   The pharmaceutical composition of claim 62 wherein the reagent is an
25   antibody.

66.   A pharmaceutical composition, comprising:
        an expression vector encoding a polypeptide comprising the amino acid
        sequence shown in SEQ ID NOS: 2 or 7; and
30   a pharmaceutically acceptable carrier.

67.   The pharmaceutical composition of claim 66 wherein the expression vector comprises SEQ ID NO: 1 or 6.

68.   A method of treating a membrane-type serine protease dysfunction related disease, wherein the disease is selected from cancer, neurodegenerative disease, osteoporosis or chronic obstructive pulmonary disease, comprising the step of:
        administering to a patient in need thereof a therapeutically effective dose of a reagent that modulates a function of a human membrane-type serine protease, whereby symptoms of the membrane-type serine protease dysfunction related disease are ameliorated.

69.   The method of claim 68 wherein the reagent is identified by the method of claim 36.

70.   The method of claim 68 wherein the reagent is identified by the method of claim 45.

71.   The method of claim 68 wherein the reagent is identified by the method of claim 49.

Fig. 1

```
gcccagcgtt   gtgatggagt   aaatgactgc   tttgatgaaa
gtgatgaact   gttttgcgtg   agccctcaac   ctgcctgcaa
taccagctcc   ttcaggcagc   atggccctct   catctgtgat
ggcttcaggg   actgtgagaa   tggccgggat   gagcaaaact
gcactcaaag   tattccatgc   aacaacagaa   cttttaagtg
tggcaatgat   atttgcttta   ggaaacaaaa   tgcaaaatgt
gatgggacag   tggattgtcc   agatggaagt   gatgaagaag
gctgcacctg   cagcaggagt   tcctccgccc   ttcaccgcat
catcggaggc   acagacaccc   tggagggggg   ttggccgtgg
caggtcagcc   tccactttgt   tggatctgcc   tactgtggtg
cctcagtcat   ctccagggag   tggcttcttt   ctgcagccca
ctgttttcat   ggaaacaggc   tgtcagatcc   cacaccatgg
actgcacacc   tcgggatgta   tgttcagggg   aatgccaagt
ttgtctcccc   ggtgagaaga   attgtggtcc   acgagtacta
taacagtcag   acttttgatt   atgatattgc   tttgctacag
ctcagtattg   cctggcctga   gaccctgaaa   cagctcattc
agccaatatg   cattcctccc   actggtcaga   gagttcgcag
tggggagaag   tgctgggtaa   ctggctgggg   gcgaagacac
gaagcagata   ataaaggctc   cctcgttctg   cagcaagcgg
aggtagagct   cattgatcaa   acgctctgtg   tttccaccta
cgggatcatc   acttctcgga   tgctctgtgc   aggcataatg
tcaggcaaga   gagatgcctg   caaaggagat   tcgggtggac
ctttatcttg   tcgaagaaaa   agtgatggaa   aatggatttt
gactggcatt   gttagctggg   gacatggatg   tggacgacca
aactttcctg   gtgtttacac   aagggtgtca   aactttgttc
cctggattca   taaatatgtc   ccttctcttt   tgtaa
```

Fig. 2

```
AQRCDGVNDC   FDESDELFCV   SPQPACNTSS   FRQHGPLICD
GFRDCENGRD   EQNCTQSIPC   NNRTFKCGND   ICFRKQNAKC
DGTVDCPDGS   DEEGCTCSRS   SSALHRIIGG   TDTLEGGWPW
QVSLHFVGSA   YCGASVISRE   WLLSAAHCFH   GNRLSDPTPW
TAHLGMYVQG   NAKFVSPVRR   IVVHEYYNSQ   TFDYDIALLQ
LSIAWPETLK   QLIQPICIPP   TGQRVRSGEK   CWVTGWGRRH
EADNKGSLVL   QQAEVELIDQ   TLCVSTYGII   TSRMLCAGIM
SGKRDACKGD   SGGPLSCRRK   SDGKWILTGI   VSWGHGCGRP
NFPGVYTRVS   NFVPWIHKYV   PSLL
```

Fig. 3

```
MGSNRGRKAG   GGSQDFGAGL   KYDSRLENMN   GFEEGVEFLP
ANNAKKVEKR   GPRRWVVLVA   VLFSFLLLSL   MAGLLVWHFH
YRNVRVQKVF   NGHLRITNEI   FLDAYENSTS   TEFISLASQV
KEALKLLYNE   VPVLGPYHKK   SAVTAFSEGS   VIAYYWSEFS
IPPHLAEEVD   RAMAVERVVT   LPPRARALKS   FVLTSVVAFP
IDPRMLQRTQ   DNSCSFALHA   HGAAVTRFTT   PGFPNSPYPA
HARCQWVLRG   DADSVLSLTF   RSFDVAPCDE   HGSDLVTVYD
SLSPMEPHAV   VRLCGTFSPS   YNLTFLSSQN   VFLVTLITNT
GRRHLGFEAT   FFQLPKMSSC   GGVLSDTQGT   FSSPYYPGHY
PPNINCTWNI   KVPNNRNVKV   RFKLFYLVDP   NVPVGSCTKD
YVEINGEKGS   GERSQFVVSS   NSSKITVHFH   SDHSYTDTGF
LAEYLSYDSN   DPCPGMFMCK   TGRCIRKELR   CDGWADCPDY
SDERYCRCNA   THQFTCKNQF   CKPLFWVCDS   VNDCGDGSDE
EGCSCPAGSF   KCSNGKCLPQ   SQKCNGKDNC   GDGSDEASCD
SVNVVSCTKY   TYRCQNGLCL   SKGNPECDGK   TDCSDGSDEK
NCDCGLRSFT   KQARVVGGTN   ADEGEWPWQV   SLHALGQGHL
CGASLISPDW   LVSAAHCFQD   DKNFKYSDYT   MWTAFLGLLD
QSKRSASGVQ   ELKLKRIITH   PSFNDFTFDY   DIALLELEKS
VEYSTVVRPI   CLPDATHVFP   AGKAIWVTGW   GHTKEGGTGA
LILQKGEIRV   INQTTCEDLM   PQQITPRMMC   VGFLSGGVDS
CQGDSGGPLS   SAEKDGRMFQ   AGVVSWGEGC   AQRNKPGVYT
RLPCSSGLDQ   RAHWGIAAWT   DSRPQTPTGM   PDMHTWIQER
NTDDIYAVAS   PPQHNPDCEL
```

Fig. 4

vrVSlsvrlGehnlsltegt.eqkfdvkktiivHpnynpdtldngaYdnDi
ALlkLksp.

gvtlgdtvrpicLpsassdlpvGttctvsGwGrrptknlg.lsdtLqevvv
pvvsretCr

sayeyggtdDkvefvtdnmiCagal.ggkdaCqGDSGGPLvcsdgnrdgrw
elvGivSwG

sygCargnkPGvytrVssyldWI

Fig. 5

stCggpdeFqCgsgrrCIprswvCDGdpDCeDGSDEslenCaa

Fig. 6

```
atggacaaag     aaaacagcga     tgtttcagcc     gcacctgctg
acctgaaaaa     gctgccagtg     agacgaccac     cgttgccagg
gagacgacta     ccattgccag     gaagacgacc     accacaaaga
cccattggca     aagccaaacc     caagaagcaa     tccaagaaaa
aagttccctt     ttggaatgta     caaaataaaa     tcattctctt
cacagtattt     ttattcatcc     tagcagtcat     agcctggaca
cttctgtggc     tgtatatcag     taagacagaa     agcaaagatg
cttttactt      tgctgggatg     tttcgcatca     ccaacatcga
gtttcttccc     gaataccgac     aaaaggagtc     cagggaattt
ctttcagtgt     cacggactgt     gcagcaagtg     ataaacctgg
tttatacaac     atctgccttc     tccaaatttt     atgagcagtc
tgttgttgca     gatgtcagca     gcaacaacaa     aggcggcctc
cttgtccact     tttggattgt     ttttgtcatg     ccacgtgcca
aaggccacat     cttctgtgaa     gactgtgttg     ccgccatctt
gaaggactcc     atccagacaa     gcatcataaa     ccggacctct
gtggggagct     tgcaggact      ggctgtggac     atggactctg
tggtactaaa     tggtgattgt     tggtcattcc     taaaaaaaa
gaaaagaaag     gaaaatggtg     ctgtctccac     agacaaaggc
tgctctcagt     acttctatgc     agagcatctg     tctctccact
acccgctgga     gatttctgca     gcctcaggga     ggctgatgtg
tcacttcaag     ctggtggcca     tagtgggcta     cctgattcgt
ctctcaatca     agtccatcca     aatcgaagcc     gacaactgtg
tcactgactc     cctgaccatt     tacgactccc     ttttgcccat
ccggagcagc     atcttgtaca     gaatttgtga     acccacaaga
acattaatgt     catttgtttc     tacaaataat     ctcatgttgg
tgacatttaa     gtctcctcat     atacggaggc     tctcaggaat
ccgggcatat     tttgagtgtg     aaaacacagt     gttggtcaaa
gacatcactg     gctttgaagg     gaaaatttca     agcccatatt
acccgagcta     ctatcctcca     aaatgcaagt     gtacctggaa
atttcagatc     atcagacaat     tttttgagtg     cccagccctc
tggttcacat     tcagctccag     tgcagttcaa     ggctttcaga
caagccactt     ttggcagaat     atggcagtta     caacatcagt
caacgaagag     ctgtttggta     gagggagttc     aaatatcagc
cctcagaata     ttgagtttga     aattcatact     gaaaactact
ggaagagcgg     ccaggaagtg     ctcatcagaa     aaggcaatgt
gatgaatgat     tcaagatcgg     gaatgaatct     tgaatcaggg
```

Fig. 6 (continued)

```
aatggatcat      taatcagaag      cctggttgaa      gaggaaagac
tagtggacag      ggaggcccag      cgttgtgatg      gagtaaatga
ctgctttgat      gaaagtgatg      aactgttttg      cggtggactg
gcactgctga      tcatcagaga      acccagctgg      tgtctcaggc
tcgtgaactt      tctgttttta      tttttattc       tagtgagccc
tcaacctgcc      tgcaatacca      gctccttcag      gcagcatggc
cctctcatct      gtgatggctt      cagggactgt      gagaatggcc
gggatgagca      aaactgcact      caaagtattc      catgcaacaa
cagaactttt      aagtgtggca      atgatatttg      ctttaggaaa
caaaatgcaa      aatgtgatgg      gacagtggat      tgtccagatg
gaagtgatga      agaaggctgc      acctgcagca      ggagttcctc
cgcccttcac      cgcatcatcg      gaggcacaga      caccctggag
gggggttggc      cgtggcaggt      cagcctccac      tttgttggat
ctgcctactg      tggtgcctca      gtcatctcca      gggagtggct
tctttctgca      gcccactgtt      ttcatggaaa      caggctgtca
gatcccacac      catggactgc      acacctcggg      atgtatgttc
aggggaatgc      caagtttgtc      tccccggtga      gaagaattgt
ggtccacgag      tactataaca      gtcagacttt      tgattatgat
attgctttgc      tacagctcag      tattgcctgg      cctgagaccc
tgaaacagct      cattcagcca      atatgcattc      ctcccactgg
tcagagagtt      cgcagtgggg      agaagtgctg      ggtaactggc
tggggcgaa       gacacgaagc      agataataaa      ggctccctcg
ttctgcagca      agcggaggta      gagctcattg      atcaaacgct
ctgtgtttcc      acctacggga      tcatcacttc      tcggatgctc
tgtgcaggca      taatgtcagg      caagagagat      gcctgcaaag
gagattcggg      tggacctta       tcttgtcgaa      gaaaaagtga
tggaaaatgg      attttgactg      gcattgttag      ctggggacat
ggatgtggac      gaccaaactt      tcctggtgtt      tacacaaggg
tgtcaaactt      tgttccctgg      attcataaat      atgtcccttc
tcttttgtaa
```

Fig. 7

```
MDKENSDVSA   APADLKKLPV   RRPPLPGRRL   PLPGRRPPQR
PIGKAKPKKQ   SKKKVPFWNV   QNKIILFTVF   LFILAVIAWT
LLWLYISKTE   SKDAFYFAGM   FRITNIEFLP   EYRQKESREF
LSVSRTVQQV   INLVYTTSAF   SKFYEQSVVA   DVSSNNKGGL
LVHFWIVFVM   PRAKGHIFCE   DCVAAILKDS   IQTSIINRTS
VGSLQGLAVD   MDSVVLNGDC   WSFLKKKKRK   ENGAVSTDKG
CSQYFYAEHL   SLHYPLEISA   ASGRLMCHFK   LVAIVGYLIR
LSIKSIQIEA   DNCVTDSLTI   YDSLLPIRSS   ILYRICEPTR
TLMSFVSTNN   LMLVTFKSPH   IRRLSGIRAY   FECENTVLVK
DITGFEGKIS   SPYYPSYYPP   KCKCTWKFQI   IRQFFECPAL
WFTFSSSAVQ   GFQTSHFWQN   MAVTTSVNEE   LFGRGSSNIS
PQNIEFEIHT   ENYWKSGQEV   LIRKGNVMND   SRSGMNLESG
NGSLIRSLVE   EERLVDREAQ   RCDGVNDCFD   ESDELFCGGL
ALLIIREPSW   CLRLVNFLFL   FFILVSPQPA   CNTSSFRQHG
PLICDGFRDC   ENGRDEQNCT   QSIPCNNRTF   KCGNDICFRK
QNAKCDGTVD   CPDGSDEEGC   TCSRSSSALH   RIIGGTDTLE
GGWPWQVSLH   FVGSAYCGAS   VISREWLLSA   AHCFHGNRLS
DPTPWTAHLG   MYVQGNAKFV   SPVRRIVVHE   YYNSQTFDYD
IALLQLSIAW   PETLKQLIQP   ICIPPTGQRV   RSGEKCWVTG
WGRRHEADNK   GSLVLQQAEV   ELIDQTLCVS   TYGIITSRML
CAGIMSGKRD   ACKGDSGGPL   SCRRKSDGKW   ILTGIVSWGH
GCGRPNFPGV   YTRVSNFVPW   IHKYVPSLL
```

FIG. 8

BLASTP - Querry = 154_extC_TR1; Hit = swiss|P56677|EPIT_MOUSE
This hit is scoring at : 5e-75 (expectation value)
Alignment length (overlap) : 344
Identities : 43 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb

Q: 4 CDGVNDCFDESDELFCVSPQPACNTSSFR-QHGPLI-----CDGFRDCENGRDEQNC--T
           CD.VNDC D SDE C   :C   .SF: .:G.:   C:G   :C :G.DE.:C .
H: 508 CDSVNDCGDGSDEEGC-----SCPAGSFKCSNGKCLPQSQKCNGKDNCGDGSDEASCDSV
           QSIPCNNRTFKCGNDICFRKQNAKCDGTVDCPDGSDEEGCTCS-RSSSALHRIIGGTDTL
           .:C.. T::C N..C..K N.:CDG..DC.DGSDE:.C.C. RS :. R::GGT:.
           NVVSCTKYTYRCQNGLCLSKGNPECDGKTDCSDGSDEKNCDCGLRSFTKQARVVGGTNAD

           EGGWPWQVSLHFVGSAY-CGASVISREWLLSAAHCFHGN---RLSDPTPWTAHLGMYVQG
           EG WPWQVSLH :G.: CGAS:IS :WL:SAAHCF.:   .:SD T WTA.LG:. Q.
           EGEWPWQVSLHALGQGHLCGASLISPDWLVSAAHCFQDDKNFKYSDYTMWTAFLGLLDQS
                                   TRYPSIN HIS

           NAKFVS-----PVRRIVHEYYNSQTFDYDIALLQLSIAWPETLKQLIQPICIPPTGQRVR
           ... .::RI:.H. :N. TFDYDIALL:L. :   . .:::::PIC:P.. .
           KRSASGVQELKLKRIITHPSFNDFTFDYDIALLELEKS--VEYSTVVRPICLPDATHVFP
                                         active site ASP

           SGEKCWVTGWGRRHEADNKGSIVLQQAEVELIDQTLCVSTY-GIITSRMLCAGIMSGKRD
           :G:..WVTGWG...E. .G:L:LQ:.E::.I:QT.C ...   IT.RM:C.G.:SG D
           AGKAIWVTGWGHTKEG-GTGALILQKGEIRVINQTTCEDLMPQQITPRMCVGFLSGGVD

           ACKGDSGGPLSCRRKSDGKWILTGIVSWGHGCGRPNFPGVYTRV   329
           :C:GDSGGPLS...K DG:.. .G:VSWG.GC.: N PGVYTR:
           SCQGDSGGPLSSAEK-DGRMFQAGVVSWGEGCAQRNKPGVYTRL   842
                     TRYPSIN SER

BLOCKS are underlinded

FIG. 9

BLOCKS search results

AC#	Description	Strength	Score
BL01253H	Type I fibronectin domain proteins.	1765	1514
AA#	304 wiLTGIvSWGhGCGRpNfPGVYTRVsNfVpWIHky (SEQ ID NO:4)		
BL01253G	Type I fibronectin domain proteins.	1641	1468
AA#	283 rDACkGDSGGPLsC (SEQ ID NO:5)		
BL00021D	Kringle domain proteins.	1556	1410
AA#	294 lscrrksDGKWILTGIVSWGhGCGRPNfPGVYTRVSNFVPWI (SEQ ID NO:6)		
BL00134A	Serine proteases, trypsin family, histidine p	1500	1409
AA#	131 CGASVISREWLLSAAHC (SEQ ID NO:7)		
BL00021B	Kringle domain proteins.	1547	1402
AA#	131 CGASvISrEWllSAAHCF (SEQ ID NO:8)		
BL01209	LDL-receptor class A (LDLRA) domain proteins.	1413	1399
AA#	79 CDGTVDCPDGSDE (SEQ ID NO:9)		
BL00134C	Serine proteases, trypsin family, histidine p	1245	1322
AA#	322 PGVYTRVSNFVPWI (SEQ ID NO:10)		

FIG. 9 (continued)

BL004950   Apple domain proteins.

1756   1277

AA#   313 GhGCGRpnfPGVYTrVSnfVpWIhkyvpS
            (SEQ ID NO:11)

BL00495N   Apple domain proteins.

1945   1268

AA#   276 AGimsGkrDACKGDSGGPLSCrrksdgkwiltgiv
            (SEQ ID NO:12)

BL00495M   Apple domain proteins.

1943   1249

AA#   206 etlkqliqPICiPptGqRvrsgekCWVTGWGrRhE
            (SEQ ID NO:13)

FIG. 10

```
HMMPFAM - Querry = 154_extC_TR1; Hit pfam|hmm|trypsin
    Trypsin
    This hit is scoring at : 258.6; Expect = 2.7e-81
    Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Q:  107 IGGTDTLEGGW--PWQVSLHFVG-----SAYCGASVISREWLLSAAHCFHG---NRLSD
        I:GG.:..G.:   PWQVSL..  .        :CG.S:IS..W:L:AAHC..G    .S.
H:    1 IvGGreaqpgsfgsPwqvslqvrsgggsrkhfCGGsLisenwVLTAAHCvsgaasapass

        PTP---WTA-HLGMYVQGNaKFVSPVRR-IVVHEYYNSQT-----FDYDIALLQLSIAw
          .:H .::G. : .V:: I:VH. YN..T    :D DIALL:L. .
        vrVSlsvrlGehnlsltegt.eqkfdvkktivHpnynpdtldngaYdnDiAllklksp.

        PETLKQLIQPICIPPTGQRVRSGEKCWVTGWGRR-HEADNkGSLVLQQAEVELIDQTLCV
        TL ..::PIC:P....:  G..C V:GWGRR  :  S .LQ:. V.::..:.C
        gvtlgdtvrpiclpsassdlpvGttctvsGwGrrptknlg.lsdtLqevvvpvvsretCr

        ST--YG-------IITSRMLCAGIMsGKRDACKGDSGGPLSCRRK-SDGKWILTGIVSWG
        S. YG    :T.M:CAG.:  G :DAC:GDSGGPL C.  .DG:W L.GIVSWG
        sayeyggtdDkvefvtdnmiCagal.ggkdaCgGDSGGPLvcsdgnrdgrwelvGivSwG

        -HGCGRPNFPGVYTRVSNFVPWI       336
        :GC.R N PGVYTRVS:::.WI
        sygCargnkPGvytrVssyldWI       259
```

FIG.11

HMMPFAM - alignment of 154_extC_TR1 against pfam|hmm|1d1_recept_a
Low-density lipoprotein receptor domain
This hit is scoring at : 9.0; Expect = 0.096
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Q: 24 PACN-TSSFR-QHGP------LICDGFRDCENGRDE--QNCTQ   56
            ..C. ...F: .G   :CDG DCE:G.DE :NC..

H: 1 stCggpdeFqCgsgrrCIprswvCDGdpDCeDGSDEslenCaa   43

FIG. 12

BLASTP - alignment of 154_v4_TR1 against swissnew|P56677|ST14_MOUSE
SUPRESSOR OF TUMORIGENICITY 14 (EC 3.4.21.-)
(EPITHIN).//:swiss|P56677|EPIT_MOUSE EPITHIN (EC 3.4.21.-).
This hit is scoring at : 5e-105 (expectation value)
Alignment length (overlap) : 839
Identities : 32 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)

Database searched : nrdb_1_;

Q: 58 AKPKKQSKKKVPFWNVQNKIILFTVFLFILAVIAWTLLWLYISKTESKDAFYFAGMFRIT
         A. .K: :K: P   V   .IF:..L L:::A .L:W : : ... F G .RIT
H: 41 ANNAKKVEKRGPRRWVLVAVLFSFLL--LSLMAGLLVWHFHYRNVRVQKV-FNGHLRIT

         NIEFLPEYRQKESREFLSVSRTVQQVINLVYT-TSAFSKFYEQSVVADVSSNNKGGLLVH
         N FL..Y...S.EF:S:::V:::L:Y. .....:::S.V. .S   :G.::::
         NEIFLDAYENSTSTEFISLASQVKEALKLLYNEVPVLGPYHKKSAVTAFS---EGSVIAY

         FWIVFVMPRAKGHIFCEDCVAAILKDSIQTSIINRTSVGSLQGLAVDMDSVVLNAGLRSD
         :W F :P H: .E:. .A: .: T   L A .:S.VL.:: .
         YWSEFSIPP---HL-AEEVDRAMAVERVVT-----LPPRARALKSFVLTSVVAFP

         YSSTI---GSDKGCSQYFYAEHLSLH-----YPLEISAASGRLMCHFKLVAIVGYLIRL
         ...: .D.CS .A. .::   :P .A.R C. L. .::L
         IDPRMLQRTQDNSCSFALHAHGAAVTRFTTPGFPNSPYPAHAR--CQWVLRGDADSVLSL

         SIKSIQIE-ADNCVTDSLTIYDSLPIRSSILYRIC--EPTRTLMSFVSTNNLMLVTFK
         .:S..: .D .:D :T:YDSL P:...:R:C   P:. L :F:S: N:.LVT.
         TFRSFDVAPCDEHGSDLVTVYDSLSPMEPHAVVRLCGTFSPSYNL-TFLSSQNVFLVTLI

         SPHIRRLSGIRAYFEVIPEQKCENTVLVKDITGFEGKISSPYYPSYYPSPYYPPKCKCTWKFTK
         : RR G..A F :P:... VL   : :G..SSPYYP.:YPP..CTW.::..
         TNTGRRHLGFEATFFQLPKMSSSCGGVL----SDTQGTFSSPYYPGHYPPNINCTWNIKVP

FIG. 12 (continued)

```
LS-TLGIALKFYNYSITKKSMKGCEHGWWEINEHMYCGSYMDHQTIFRVPSPLVHIQLQC
: .: :..K.: . : .C.. : EIN   GS   .Q . . S. : .:....:
NNRNVKVRFKLFYLVDPNVPVGSCTKDYVEINGEK--GSGERSQFVVSSNSSKITVHFHS

SSRLSDKPLLAEYGSYNISQPCPVGSFRCSSGLCVPQAQRCDGVNDCFDESDELFC---V
: .: :..D. .LAEY SY: :.PCP G.F.C.:G C: :. RCDG  DC D SDE :C .
....:D. .LAEY.LSYDSNDPCP-GMFMCKTGRCIRKELRCDGWADCPDYSDERYCRCNA

                        LDLRA_1

weakly similar to pfam|hmm|ldl recept a

SPQPACNTSSFRQHGPL--ICDGFRDCENGRDEQNCT----------------QSIPCNNR
: .Q. .C.. :    PL   :CD...DC :G.DE::.C:            QS  .CN.::
THQFTCKNQFCK---PLFWVCDSVNDCGDGSDEEGCSCPAGSFKCSNGKCLPQSQKCNGK

                        LDLRA
--------------------------TFKCGNDICFRKQNAKCDGTVDCPDGSDEEGCTCS-R
                          T::C N.:C.. K N.:CDG..DC.DGSDE:.C.C. R
DNCGDGSDEASCDSVNVVSCTKYTYRCQNGLCLSKGNPECDGKTDCSDGSDEKNCDCGLR
                        trypsin kringle domain
                Type I fibronectin Apple domain
SSSALHRIIGGTDTLEGGWPWQVSLHFVGSAY-CGASVISREWLLSAAHCFHGN---RLS
S :. R::GGT:. EG WPWQVSLH :G..: CGAS:IS :WL:SAAHCF...: ..:S
SFTKQARVVGGTNADEGEWPWQVSLHALGQGHLCGASLISPDWLVSAAHCFQDDKNFKYS
                        TRYPSIN_HIS
DPTPWTAHLGMYVQGNAKFVS-----PVRRIVVHEYNSQTFDYDIALLQLSIAWPETLKQ
D T WTA.LG:. Q....  ..:RI::H. :N. TFDYDIALL:L. :  .:...:
DYTMWTAFLGLLDQSKRSASGVQELKLKRIITHPSFNDFTFDYDIALLELEKS--VEYST
                        ACT_SITE_ASP
```

FIG. 12 (continued)

Type I fibronectin Apple domain

```
LIQPICIPPTGQRVRSGEKCWVTGWGRRHEADNKGSLVLQQAEVELIDQTLCVSTY-GII
:::PIC:P.. . :G:..WVTGWG...E. ..G:L:LQ:.E:.:I:QT.C ... I
VVRPICLPDATHVFPAGKAIWVTGWGHTKEG-GTGALILQKGEIRVINQTTCEDLMPQQI
```

Apple domain Kringle domain trypsin Type I fibronectin

```
TSRMLCAGIMSGKRDACKGDSGGPLSCRRKSDGKWILTGIVSWGHGSGRPNFPGVYTRV  828
T.RM:C.G.:SG   D:C:GDSGGPLS...K DG:...  .G:VSWG.G..: N PGVYTR:
TPRMCVGFLSGGVDSCQGDSGGPLSSAEK-DGRMFQAGVVSWGEGCAQRNKPGVYTRL   842
```

TRYPSIN_SER

BLOCKS are underlined
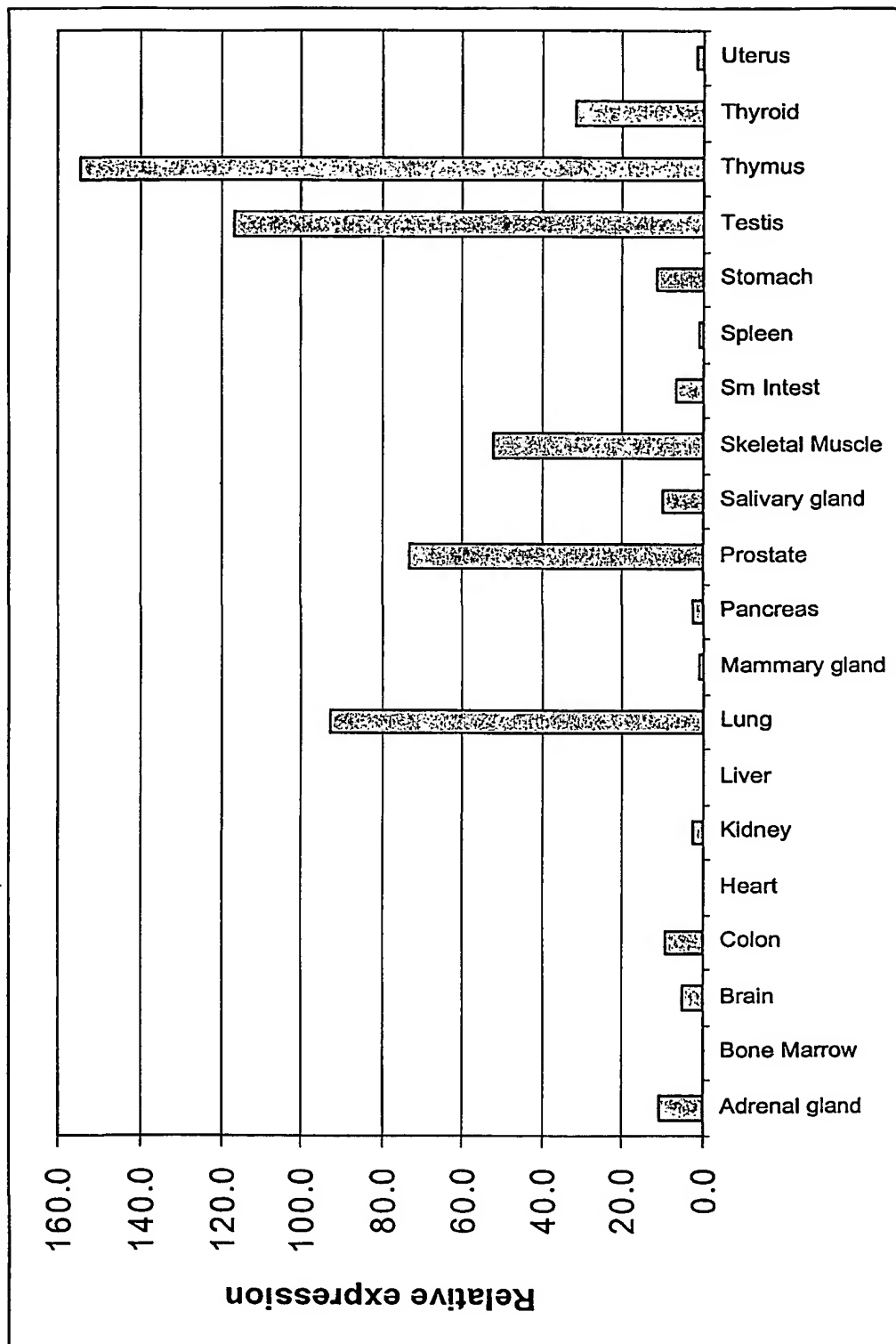
Fig. 13. Relative expression of membrane-type serine protease in various human tissues.

Fig. 14. Relative expression of membrane-type serine protease in various human respiratory tissues and cells.
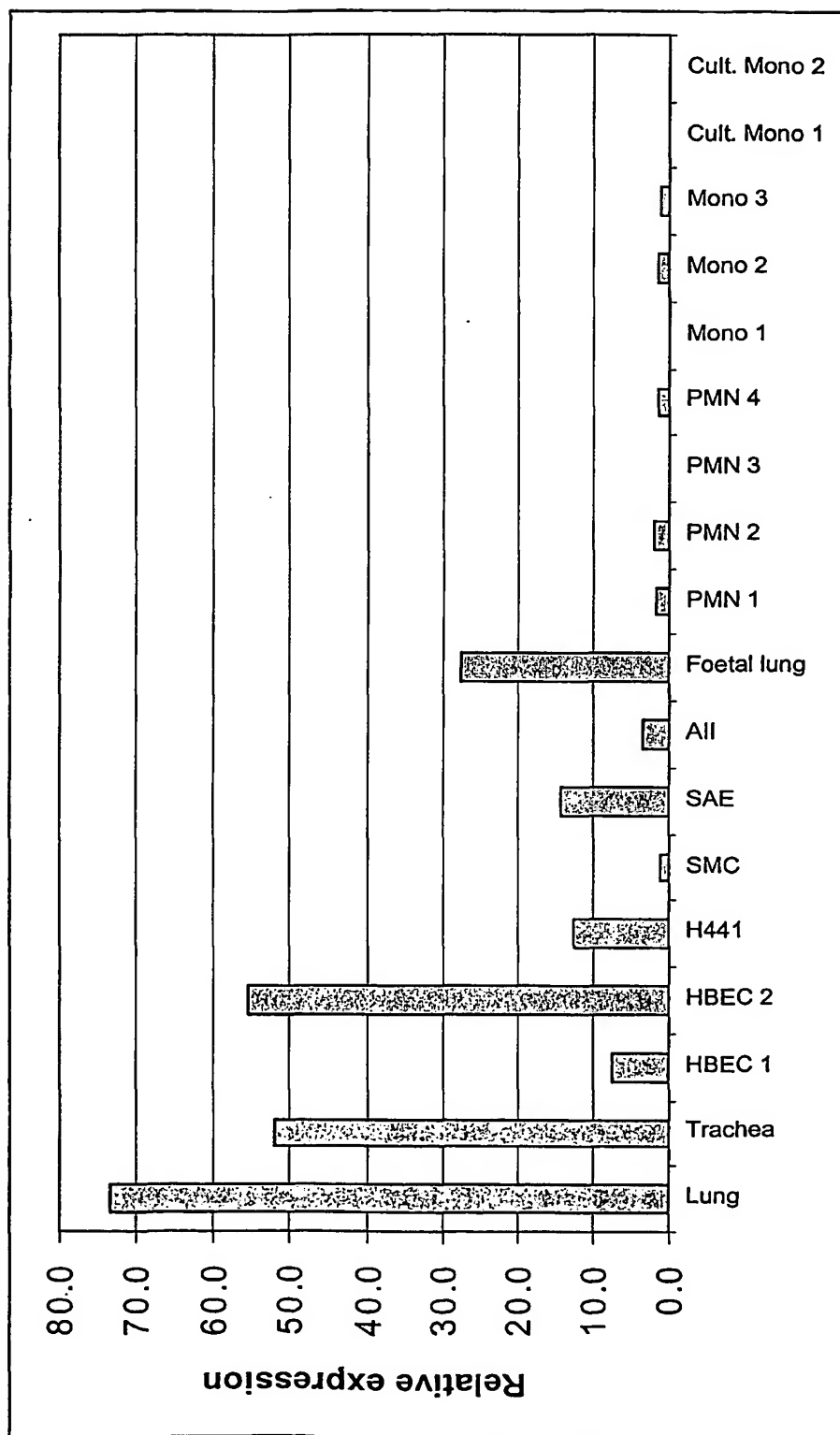
Key: HBEC=cultured human bronchial epithelial cells; H441=Clara-like cells; SMC=cultured airway smooth muscle cells; SAE= cultured small airway epithelial cells; AII=primary cultured alveolar type II cells; PMN=polymorphonuclear leukocytes; Mono=monocytes; Cult. Mono=cultured monocytes (macrophage-like).

Fig 15

Prosite search results

Accession#	From->To	Name	Doc#
PS01209	496->519	LDLRA_1	PDOC00929
PS00134	642->648	TRYPSIN_HIS	PDOC00124
PS00135	784->796	TRYPSIN_SER	PDOC00124

Fig. 16 BLOCKS search results

AC#	Description		Strength	Score
BL01253G	Type I fibronectin domain proteins.		1453	1586
AA#	782	rDACkGDSGGPLsC		
BL00021B	Kringle domain proteins.		1466	1462
AA#	630	CGASvISrEWLlSAAHCF		
BL01209	LDL-receptor class A (LDLRA) domain proteins.		1296	1428
AA#	578	CDGTVDCPDGSDE		
BL00134A	Serine proteases, trypsin family, histidine p	1439	1425	
AA#	630	CGASVISREWLLSAAHC		
BL00495N	Apple domain proteins.		1538	1319
AA#	775	AGimsGkrDACKGDSGGGPLSCrrksdgkwiltgiv		
BL00495M	Apple domain proteins.		1493	1275
AA#	705	etlkqliqPICiPptGqRvrsgekCWVTGWGrRhE		

Fig. 16 (continued)

BL01253H   Type I fibronectin domain proteins.   1520   1230

AA#   803 wiLTGIvSWGhGsGRpNfPGVYTRV

BL00134B   Serine proteases, trypsin family,   1225   1222
            histidine p

AA#   783 DACKGDSGGPLSCRRKSDGkWILT

BL01253E   Type I fibronectin domain proteins.   1385   1220

AA#   703 wpeTLkQliQPICiPPtGqRvrsGeKCwVtGWGrrhE

BL00021D   Kringle domain proteins.   1410   1212

AA#   793 lscrrksDGKWILTGIVSWGhGsGRPNfPGVYTRV

BL01253D   Type I fibronectin domain proteins.   1257   1205

AA#   630 CGaSvISreWllSA

BL00495K   Apple domain proteins.   1438   1203

AA#   632 aSvIsReWllsAAHCFhGnrLsDptpwtahlgm

Fig.17

HMMPFAM - alignment of 154_v4_TR1 against pfam|hmm|trypsin

Trypsin

This hit is scoring at : 230.9; Expect = 1.7e-72

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:    606  IIGGTDTLEGGW--PWQVSLHFVG-----SAYCGASVISREWLLSAAHCFHG---NRLSD
           I:GG.:. .G.:  PWQVSL..  .    :CG.S:IS..W:L:AAHC..G   .S.
H:      1  IvGGreaqpgsfgsPwqvslqvrsgggsrkhfCGGsLisenwVLTAAHCvsgaasapass

      PTP----WTA-HLGMYVQGNaKFVSPVRR-IVVHEYYNSQT----FDYDIALLQLSIAw
      . H  ...:G. . V:: I:VH. YN..T    :D DIALL:L. .
      vrVSlsvrlGehnlsltegt.eqkfdvkktiivHpnynpdtldngaYdnDiAllkLksp.

      PETLKQLIQPICIPPTGQRVRSGEKCWVTGWGRR-HEADNkGSlVLQQAEVELIDQTLCV
      TL .:::PIC:P.... G.C V:GWGRR  :  . S .LQ:. V::::..C
      gvtlgdtvrpiclpsassdlpvGttctvsGwGrrptknlg.lsdtlqevvpvvsretCr

      ST--YG------IITSRMLCAGIMsGKRDACKGDSGGGPLSCRRK-SDGKWILTGIVSWG
      S. YG     .:T..M:CAG.:  G :DAC:GDSGGPL C.  .DG:W L.GIVSWG
      sayeyggtdDkvefvtdnmiCagal.ggkdaCgGDSGGPLvcsdgnrdgrwelvGivSwG

      -HGSGRPNFPGVYTRV      828
      :G..R N PGVYTRV
      sygCargnkPGvytrV      252
```

Fig. 18

HMMPFAM - alignment of 154_v4_TR1 against pfam|hmm|ldl_recept_a

Low-density lipoprotein receptor domain

This hit is scoring at : 48.0; Expect = 2.1e-10

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:  482 QPCP-VGSFRCSSGL-CVPQAQRCDGVNDCFDESDE--LFCVS  520
            ..C    .F:C.SG  C:P::  CDG  DC D SDE    C.:
H:    1 stCggpdeFqCgsgrrCIprswvCDGdpDCeDGSDEslenCaa   43
```

Fig. 19

HMMPFAM - alignment of 154_v4_TR1 against pfam|hmm|ldl_recept_a

Low-density lipoprotein receptor domain

This hit is scoring at : 9.0; Expect = 0.096

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:  523 PACN-TSSFR-QHGP------LICDGFRDCENGRDE--QNCTQ    555
            ..C. ...F:  .G        :CDG   DCE:G.DE  :NC..
H:    1 stCggpdeFqCgsgrrCIprswvCDGdpDCeDGSDEslenCaa     43
```

Fig. 20

HMMPFAM - alignment of 154_v4_TR1 against pfam|hmm|ld1_recept_a

Low-density lipoprotein receptor domain

This hit is scoring at : 17.3; Expect = 0.013

Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:  557 IPCN-NRTFKCGNDI-CFRKqNAKCDGTVDCPDGSDE--EGCTC   596
            .C.  .F:CG:.  C. : :  CDG. DC.DGSDE  E.C..
H:    1 stCggpdeFqCgsgrrCIpr.swvCDGdpDCeDGSDEslenCaa   43
```

Fig. 21   Exon-intron structure of human membrane-type serine protease

```
exon 1
Q:      1 atggacaagagaaaacagcagcgatgtttcagccgcacctgctgacctgaaa        48
H: 198457 atggacaagagaaaacagcagcgatgtttcagccgcacctgctgacctgaaa(gta  198504

exon 2
Q:     49          atatccaatatctccagtccaagtggtcagtgcccaaaagaagctgccagtgagac
H: 201684 tgaag)atatccaatatctccagtccaagtggtcagtgcccaaaagaagctgccagtgagac

          gaccaccgttgccaggaggacgactaccattgccaggaagacgaccaccaaaagacccca
          gaccaccgttgccaggaggacgactaccattgccaggaagacgaccaccacaaagacccca

          ttggcaaagccaaacccaagaagcaatccaagaaaaaagttcccttttggaatgtacaaa
          ttggcaaagccaaacccaagaagcaatccaagaaaaaagttcccttttggaatgtacaaa

          ataaaatcattctcttcacagtattttattcatcctagcagtcatagcctggacacttc
          ataaaatcattctcttcacagtattttattcatcctagcagtcatagcctggacacttc

          tgtggctgtatatca          298
          tgtggctgtatatca(gtaa  201933

exon 3
Q:    299          gtaagacagaaagcaaagatgctttttacttgctgggatgtttcgcatcaccaacatcg
H: 205532 cag)gtaagacagaaagcaaagatgctttttacttgctgggatgtttcgcatcaccaacatcg

          agtttcttcccgaataccgacaaaaggagtccaggggaatttctttcagtgtcacggactg
          agtttcttcccgaataccgacaaaaggagtccaggggaatttctttcagtgtcacggactg

          tgcagcaagtg          429
          tgcagcaagtg(gtg   205662
```

Fig. 21   (continued)

exon 4

```
Q:      430    ataaacctggtttatacaacatctgcctctccaaatttatgagcagtctgttgttgc
H:   207865 cag)ataaacctggtttatacaacatctgcctctccaaatttatgagcagtctgttgttgc

               agatgtcag    497
               agatgtcag(gta  207932
```

exon 5

```
Q:      498    cagcaacaacaaaggcggcctcttgtccactttggattgtttttgtcatgccacgt
H:   209362 tag)cagcaacaacaaaggcggcctcttgtccactttggattgtttttgtcatgccacgt

               gccaaggccacatctctctgtgaagactgtgttgccgccatcttgaaggactccatccag
               gccaaggccacatctctctgtgaagactgtgttgccgccatcttgaaggactccatccag

               acaagcatcataaaccggacctctgtggggagcttgcagggactgctgtggacatggac
               acaagcatcataaaccggacctctgtggggagcttgcagggactgctgtggacatggac

               tctgtggtactaaatg    691
               tctgtggtactaaatg(gtg  209556
```

exon 6

```
Q:      692    ctgggcttcggtcagattactcgtcaaccataggatctg    730
               CTGGGCTTCGGTC GATTACTCGTCAACCATAGGATCTG

H:   210587 cag)ctgggcttcggtcagattactcgtcaaccataggatctg(gta  210625
(This exon is provided by experimental data)
```

exon 7

```
Q:      731    acaaaggctgctctcagtacttctatgcagagcatctgtctctccactaccgctggag
H:   211352 cag)acaaaggctgctctcagtacttctatgcagagcatctgtctctccactaccgctggag
```

Fig. 21   (continued)

```
atttctgcagcctcagggaggctgatgtgtgtcacttcaagctggtggtggccatagtgggctac
atttctgcagcctcagggaggctgatgtgtgtcacttcaagctggtggtggccatagtgggctac

ctgattcgtctctcaatcagtccatccaaatcgaagccgacaactgtgtcactgactcc
ctgattcgtctctcaatcagtccatccaaatcgaagccgacaactgtgtcactgactcc

ctgaccattacgactccctttgccatccggagcagcatcttgtacag          959
ctgaccattacgactccctttgccatccggagcagcatcttgtacag(gta   211580
```

exon 8
```
Q:      960    aatttgtgaacccacaagaacattaatgtcatttgtttctacaataatctcatgtt
H:   213457  cag)aatttgtgaacccacaagaacattaatgtcatttgtttctacaataatctcatgtt

ggtgacatttaagtctcctcatatacggaggctctcaggaatccgggcatatttgaggt
ggtgacatttaagtctcctcatatacggaggctctcaggaatccgggcatatttgaggt

cattccagaacaaa          1090
cattccagaacaaa(gta   213587
```

exon 9
```
Q:     1091   agtgtgaaaacacagtgttggtcaaagacatcactggctttgaagggaaatttcaagcc
H:   217373  cag)agtgtgaaaacacagtgttggtcaaagacatcactggctttgaagggaaatttcaagcc

catattacccgagctactatcctccaaaatgcaagtgtacctggaaatttcag          1203
catattacccgagctactatcctccaaaatgcaagtgtacctggaaatttcag(gta   217485
```

Fig. 21   (continued)

```
exon 10
Q:  1204    acttctctatcaactcttggcatagcactgaaattctataactattcaataaccaag
H: 223729 cag)acttctctatcaactcttggcatagcactgaaattctataactattcaataaccaag

            aagagtatgaaaggctgtgagcatggatggtggggaaattaatgagcacat    1310
            aagagtatgaaaggctgtgagcatggatggtggggaaattaatgagcacat(gta  223835

exon 11
Q:  1311    gtactgtggctcctacatggatcatcagacaatttttcgagtgccagcccctctggttca
H: 228491 cag)gtactgtggctcctacatggatcatcagacaatttttcgagtgccagcccctctggttca

            cattcagctccagtgcagttcaaggctttcagacaagccactttttggcagaatatggcag
            cattcagctccagtgcagttcaaggctttcagacaagccactttttggcagaatatggcag

            ttacaacatcagtcaac    1447
            ttacaacatcagtcaac(gta  228627

exon 12
Q:  1448    cctgccctgttggatcttttagatgctcctccggtttatgtgtccctcaggcccagcgtt
            CCTGCCCTGTTGGATCTTTTAGATGCTCCTCCGGTTTATGTGTCCCTCAGGCCCAGCCAGCG T
H: 230229 tag)cctgccctgttggatctttttagatgctcctccggtttatgtgtccctcaggcccagcg-t

            gtgatgggagtaaatgactgctttgatgaaagtgatgaa            1545
            GTGATGGAGTAAATGACTGCTTTGA GAAAG GATGAA
            gtgatgggagtaaatgactgctttgaagaaagagatgaat(gctt  230325

Alignemnt of this exon against NT_023818.2:
Q:  1448    cctgccctgttggatctttttagatgctcctccggtttatgtgtccctcaggcccagcgtt
            CCTGCCCTGTTGGATCTTTTAGATGCTCCTCCGGTTTATGTGTCCCTCAGGCCCAGCGTT
```

Fig. 21   (continued)

H: 167271 tag)cctgccctgttgatctttagatgctcctccggttatgtgtccctcaggccagcgtt

```
          gtgatggagtaaatgactgctttgatgaaagtgatgaactgttttgcg          1555
          GTGATGGAGTAAATGACTGCTTTGATGAAAGTGATGAACTGTTTTGCG
          gtgatggagtaaatgactgctttgatgaaagtgatgaactgttttgcg(gtg  167378
```

exon 13

Q: 1556   tgagccctcaacctgcctgcaataccagctccttcaggcagcagtggccctctcatctgt

H: 233134 tag)tgagccctcaacctgcctgcaataccagctcctttcaggcagcagtggccctctcatctgt

```
          gatggcttcaggggactgtgagaatggccgggatgagcaaaactgcactcaaa     1666
          gatggcttcaggggactgtgagaatggccgggatgagcaaaactgcactcaaa(gtg  233244
```

exon 14

Q: 1667   gtattccatgcaacaacagaactttaagtgtggcaatgatatttgctttaggaaacaaa

H: 241036 tag)gtattccatgcaacaacagaactttaagtgtggcaatgatatttgctttaggaaacaaa

```
          atgcaaaatgtgatgggacagtggattgtgtccagatggaagtgatgaagaggctgca     1783
          atgcaaaatgtgatgggacagtggattgtgtccagatggaagtgatgaagaggctgca(gta  241152
```

exon 15

Q: 1784   cctgcagcaggagttcctccgcgccctttcaccgcatcatcggaggcacagacaccctggagg

H: 242061 aag)cctgcagcaggagttcctccgcgccctttcaccgcatcatcggaggcacagacaccctggagg

```
          ggggttggccgtggcaggtcagcctccacttTgttggatctgcctcagctggtggtgcctcag
          ggggttggccgtggcaggtcagcctccacttTgttggatctgcctcagctggtggtgcctcag
```

Fig. 21   (continued)

```
                    tcatctccaggggagtggcttcttttctgcagccccactgtttttcatgggaaacag              1955
                    tcatctccaggggagtggcttcttttctgcagccccactgtttttcatggaaacag(gta   242232

exon 16
Q:     1956              gctgtcagatcccacaccatggactgcacacctcgggatgtatgttcaggggggaatgc
H:  243630    cag)gctgtcagatcccacaccatggactgcacacctcgggatgtatgttcaggggggaatgc

                    caagtttgtctcccgtgagaagaattgtggtccacgagtactataacagtcagacttt
                    caagtttgtctcccgtgagaagaattgtggtccacgagtactataacagtcagacttt

                    tgattatgatattgctttgctacagctcagtattgcctggctcagagtcgtggggagaagtgctg
                    tgattatgatattgctttgctacagctcagtattgcctggctcagagtcgtggggagaagtgctg

                    cattcagccaatatgcattcctcccactggtcagagagttcgcagtgggggagaagtgctg
                    cattcagccaatatgcattcctcccactggtcagagagttcgcagtgggggagaagtgctg

                    ggtaactggctgggggcgaagacacgaagcag         2224
                    ggtaactggctgggggcgaagacacgaagcag(gtg   243898

exon 17
Q:     2225              ataataaaggctccctcgttctgcagcaagcggaggtagagctcattgatcaaacg
H:  245496    gcag)ataataaaggctccctcgttctgcagcaagcggaggtagagctcattgatcaaacg

                    ctctgtgtttccacctacgggatcatcacttctcggatgctctgtgcaggcataatgtca
                    ctctgtgtttccacctacgggatcatcacttctcggatgctctgtgcaggcataatgtca

                    ggcaagagagatgcctgcaaa         2361
                    ggcaagagagatgcctgcaaa(gta   245632
```

Fig. 21   (continued)

```
exon 18
Q:   2362      ggagattcgggtggacctttatcttgtcgaagaaaagtgatggaaaatggattttgact
H: 247668 tag)ggagattcgggtggacctttatcttgtcgaagaaaaagtgatggaaaatggattttgact

          ggcattgttagctgggggacatggaagtggacgaccaaactttcctggtgtttacacaagg
          ggcattgttagctgggggacatggaagtggacgaccaaactttcctggtgtttacacaagg

          gtgtcaaactttgttccctggattcataaatatgtccctctttttgtaa    2532
          gtgtcaaacttttgttccctggattcataaaatatgtccctctttttgtaa 247838
```

SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN MEMBRANE-TYPE SERINE PROTEASE

<130> LIO130 Foreign Countries

<150> US 60/223,479
<151> 2000-08-07

<150> US 60/280,132
<151> 2001-04-02

<150> US 60/XXX,XXX
<151> 2001-06-08

<160> 7

<170> PatentIn version 3.1

<210> 1
<211> 1035
<212> DNA
<213> Homo sapiens

<400> 1
gcccagcgtt gtgatggagt aaatgactgc tttgatgaaa gtgatgaact gttttgcgtg   60

agccctcaac ctgcctgcaa taccagctcc ttcaggcagc atggccctct catctgtgat   120

ggcttcaggg actgtgagaa tggccgggat gagcaaaact gcactcaaag tattccatgc   180

aacaacagaa cttttaagtg tggcaatgat atttgcttta ggaaacaaaa tgcaaaatgt   240

gatgggacag tggattgtcc agatggaagt gatgaagaag ctgcacctg cagcaggagt   300

tcctccgccc ttcaccgcat catcggaggc acagacaccc tggagggggg ttggccgtgg   360

caggtcagcc tccactttgt tggatctgcc tactgtggtg cctcagtcat ctccagggag   420

tggcttcttt ctgcagccca ctgtttcat ggaaacaggc tgtcagatcc cacaccatgg   480

actgcacacc tcgggatgta tgttcagggg aatgccaagt ttgtctcccc ggtgagaaga   540

attgtggtcc acgagtacta taacagtcag acttttgatt atgatattgc tttgctacag   600

ctcagtattg cctggcctga gaccctgaaa cagctcattc agccaatatg cattcctccc   660

actggtcaga gagttcgcag tggggagaag tgctgggtaa ctggctgggg gcgaagacac   720

gaagcagata taaaggctc cctcgttctg cagcaagcgg aggtagagct cattgatcaa   780

acgctctgtg tttccaccta cgggatcatc acttctcgga tgctctgtgc aggcataatg   840

tcaggcaaga gagatgcctg caaaggagat tcgggtggac ctttatcttg tcgaagaaaa   900

agtgatggaa aatggatttt gactggcatt gttagctggg gacatggatg tggacgacca   960

aactttcctg gtgtttacac aagggtgtca aactttgttc cctggattca taaatatgtc   1020

ccttctcttt tgtaa   1035

```
<210>   2
<211>   344
<212>   PRT
<213>   Homo sapiens

<400>   2

Ala Gln Arg Cys Asp Gly Val Asn Asp Cys Phe Asp Glu Ser Asp Glu
1               5                   10                  15


Leu Phe Cys Val Ser Pro Gln Pro Ala Cys Asn Thr Ser Ser Phe Arg
            20                  25                  30


Gln His Gly Pro Leu Ile Cys Asp Gly Phe Arg Asp Cys Glu Asn Gly
        35                  40                  45


Arg Asp Glu Gln Asn Cys Thr Gln Ser Ile Pro Cys Asn Asn Arg Thr
    50                  55                  60


Phe Lys Cys Gly Asn Asp Ile Cys Phe Arg Lys Gln Asn Ala Lys Cys
65                  70                  75                  80


Asp Gly Thr Val Asp Cys Pro Asp Gly Ser Asp Glu Glu Gly Cys Thr
            85                  90                  95


Cys Ser Arg Ser Ser Ser Ala Leu His Arg Ile Ile Gly Gly Thr Asp
            100                 105                 110


Thr Leu Glu Gly Gly Trp Pro Trp Gln Val Ser Leu His Phe Val Gly
        115                 120                 125


Ser Ala Tyr Cys Gly Ala Ser Val Ile Ser Arg Glu Trp Leu Leu Ser
    130                 135                 140


Ala Ala His Cys Phe His Gly Asn Arg Leu Ser Asp Pro Thr Pro Trp
145                 150                 155                 160


Thr Ala His Leu Gly Met Tyr Val Gln Gly Asn Ala Lys Phe Val Ser
            165                 170                 175


Pro Val Arg Arg Ile Val Val His Glu Tyr Tyr Asn Ser Gln Thr Phe
        180                 185                 190


Asp Tyr Asp Ile Ala Leu Leu Gln Leu Ser Ile Ala Trp Pro Glu Thr
        195                 200                 205


Leu Lys Gln Leu Ile Gln Pro Ile Cys Ile Pro Pro Thr Gly Gln Arg
    210                 215                 220
```

Val Arg Ser Gly Glu Lys Cys Trp Val Thr Gly Trp Gly Arg Arg His
225   230   235   240

Glu Ala Asp Asn Lys Gly Ser Leu Val Leu Gln Gln Ala Glu Val Glu
                245   250   255

Leu Ile Asp Gln Thr Leu Cys Val Ser Thr Tyr Gly Ile Ile Thr Ser
            260   265   270

Arg Met Leu Cys Ala Gly Ile Met Ser Gly Lys Arg Asp Ala Cys Lys
        275   280   285

Gly Asp Ser Gly Gly Pro Leu Ser Cys Arg Arg Lys Ser Asp Gly Lys
    290   295   300

Trp Ile Leu Thr Gly Ile Val Ser Trp Gly His Gly Cys Gly Arg Pro
305   310   315   320

Asn Phe Pro Gly Val Tyr Thr Arg Val Ser Asn Phe Val Pro Trp Ile
                325   330   335

His Lys Tyr Val Pro Ser Leu Leu
                340

<210> 3
<211> 902
<212> PRT
<213> Mus musculus

<400> 3

Met Gly Ser Asn Arg Gly Arg Lys Ala Gly Gly Gly Ser Gln Asp Phe
1   5   10   15

Gly Ala Gly Leu Lys Tyr Asp Ser Arg Leu Glu Asn Met Asn Gly Phe
            20   25   30

Glu Glu Gly Val Glu Phe Leu Pro Ala Asn Asn Ala Lys Lys Val Glu
        35   40   45

Lys Arg Gly Pro Arg Arg Trp Val Val Leu Val Ala Val Leu Phe Ser
    50   55   60

Phe Leu Leu Leu Ser Leu Met Ala Gly Leu Leu Val Trp His Phe His
65   70   75   80

Tyr Arg Asn Val Arg Val Gln Lys Val Phe Asn Gly His Leu Arg Ile
                85   90   95

3

```
Thr Asn Glu Ile Phe Leu Asp Ala Tyr Glu Asn Ser Thr Ser Thr Glu
        100              105              110

Phe Ile Ser Leu Ala Ser Gln Val Lys Glu Ala Leu Lys Leu Leu Tyr
        115              120              125

Asn Glu Val Pro Val Leu Gly Pro Tyr His Lys Lys Ser Ala Val Thr
        130              135              140

Ala Phe Ser Glu Gly Ser Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser
145              150              155              160

Ile Pro Pro His Leu Ala Glu Glu Val Asp Arg Ala Met Ala Val Glu
                165              170              175

Arg Val Val Thr Leu Pro Pro Arg Ala Arg Ala Leu Lys Ser Phe Val
                180              185              190

Leu Thr Ser Val Val Ala Phe Pro Ile Asp Pro Arg Met Leu Gln Arg
        195              200              205

Thr Gln Asp Asn Ser Cys Ser Phe Ala Leu His Ala His Gly Ala Ala
        210              215              220

Val Thr Arg Phe Thr Thr Pro Gly Phe Pro Asn Ser Pro Tyr Pro Ala
225              230              235              240

His Ala Arg Cys Gln Trp Val Leu Arg Gly Asp Ala Asp Ser Val Leu
                245              250              255

Ser Leu Thr Phe Arg Ser Phe Asp Val Ala Pro Cys Asp Glu His Gly
        260              265              270

Ser Asp Leu Val Thr Val Tyr Asp Ser Leu Ser Pro Met Glu Pro His
        275              280              285

Ala Val Val Arg Leu Cys Gly Thr Phe Ser Pro Ser Tyr Asn Leu Thr
        290              295              300

Phe Leu Ser Ser Gln Asn Val Phe Leu Val Thr Leu Ile Thr Asn Thr
305              310              315              320

Gly Arg Arg His Leu Gly Phe Glu Ala Thr Phe Phe Gln Leu Pro Lys
                325              330              335

Met Ser Ser Cys Gly Gly Val Leu Ser Asp Thr Gln Gly Thr Phe Ser
        340              345              350
```

4

Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro Asn Ile Asn Cys Thr Trp
        355   360   365

Asn Ile Lys Val Pro Asn Asn Arg Asn Val Lys Val Arg Phe Lys Leu
    370   375   380

Phe Tyr Leu Val Asp Pro Asn Val Pro Val Gly Ser Cys Thr Lys Asp
385   390   395   400

Tyr Val Glu Ile Asn Gly Glu Lys Gly Ser Gly Glu Arg Ser Gln Phe
            405   410   415

Val Val Ser Ser Asn Ser Ser Lys Ile Thr Val His Phe His Ser Asp
        420   425   430

His Ser Tyr Thr Asp Thr Gly Phe Leu Ala Glu Tyr Leu Ser Tyr Asp
        435   440   445

Ser Asn Asp Pro Cys Pro Gly Met Phe Met Cys Lys Thr Gly Arg Cys
    450   455   460

Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp Ala Asp Cys Pro Asp Tyr
465   470   475   480

Ser Asp Glu Arg Tyr Cys Arg Cys Asn Ala Thr His Gln Phe Thr Cys
            485   490   495

Lys Asn Gln Phe Cys Lys Pro Leu Phe Trp Val Cys Asp Ser Val Asn
        500   505   510

Asp Cys Gly Asp Gly Ser Asp Glu Glu Gly Cys Ser Cys Pro Ala Gly
    515   520   525

Ser Phe Lys Cys Ser Asn Gly Lys Cys Leu Pro Gln Ser Gln Lys Cys
    530   535   540

Asn Gly Lys Asp Asn Cys Gly Asp Gly Ser Asp Glu Ala Ser Cys Asp
545   550   555   560

Ser Val Asn Val Val Ser Cys Thr Lys Tyr Thr Tyr Arg Cys Gln Asn
            565   570   575

Gly Leu Cys Leu Ser Lys Gly Asn Pro Glu Cys Asp Gly Lys Thr Asp
        580   585   590

Cys Ser Asp Gly Ser Asp Glu Lys Asn Cys Asp Cys Gly Leu Arg Ser

5

```
            595                      600                      605
```

Phe Thr Lys Gln Ala Arg Val Val Gly Gly Thr Asn Ala Asp Glu Gly
    610   615   620

Glu Trp Pro Trp Gln Val Ser Leu His Ala Leu Gly Gln Gly His Leu
625   630   635   640

Cys Gly Ala Ser Leu Ile Ser Pro Asp Trp Leu Val Ser Ala Ala His
            645   650   655

Cys Phe Gln Asp Asp Lys Asn Phe Lys Tyr Ser Asp Tyr Thr Met Trp
        660   665   670

Thr Ala Phe Leu Gly Leu Leu Asp Gln Ser Lys Arg Ser Ala Ser Gly
        675   680   685

Val Gln Glu Leu Lys Leu Lys Arg Ile Ile Thr His Pro Ser Phe Asn
    690   695   700

Asp Phe Thr Phe Asp Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Ser
705   710   715   720

Val Glu Tyr Ser Thr Val Val Arg Pro Ile Cys Leu Pro Asp Ala Thr
            725   730   735

His Val Phe Pro Ala Gly Lys Ala Ile Trp Val Thr Gly Trp Gly His
            740   745   750

Thr Lys Glu Gly Gly Thr Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile
        755   760   765

Arg Val Ile Asn Gln Thr Thr Cys Glu Asp Leu Met Pro Gln Gln Ile
    770   775   780

Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser
785   790   795   800

Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser Ser Ala Glu Lys Asp Gly
            805   810   815

Arg Met Phe Gln Ala Gly Val Val Ser Trp Gly Glu Gly Cys Ala Gln
        820   825   830

Arg Asn Lys Pro Gly Val Tyr Thr Arg Leu Pro Cys Ser Ser Gly Leu
        835   840   845

6

Asp Gln Arg Ala His Trp Gly Ile Ala Ala Trp Thr Asp Ser Arg Pro
    850   855   860

Gln Thr Pro Thr Gly Met Pro Asp Met His Thr Trp Ile Gln Glu Arg
865   870   875   880

Asn Thr Asp Asp Ile Tyr Ala Val Ala Ser Pro Pro Gln His Asn Pro
                885   .   890   895

Asp Cys Glu Leu His Pro
            900

<210> 4
<211> 199
<212> PRT
<213> Homo sapiens

<400> 4

Val Arg Val Ser Leu Ser Val Arg Leu Gly Glu His Asn Leu Ser Leu
1   5   10   15

Thr Glu Gly Thr Glu Gln Lys Phe Asp Val Lys Lys Thr Ile Ile Val
            20   25   30

His Pro Asn Tyr Asn Pro Asp Thr Leu Asp Asn Gly Ala Tyr Asp Asn
        35   40   45

Asp Ile Ala Leu Leu Lys Leu Lys Ser Pro Gly Val Thr Leu Gly Asp
    50   55   60

Thr Val Arg Pro Ile Cys Leu Pro Ser Ala Ser Ser Asp Leu Pro Val
65   70   75   80

Gly Thr Thr Cys Thr Val Ser Gly Trp Gly Arg Arg Pro Thr Lys Asn
            85   90   95

Leu Gly Leu Ser Asp Thr Leu Gln Glu Val Val Val Pro Val Val Ser
            100   105   110

Arg Glu Thr Cys Arg Ser Ala Tyr Glu Tyr Gly Gly Thr Asp Asp Lys
        115   120   125

Val Glu Phe Val Thr Asp Asn Met Ile Cys Ala Gly Ala Leu Gly Gly
    130   135   140

Lys Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Ser Asp
145   150   155   160

7

```
Gly Asn Arg Asp Gly Arg Trp Glu Leu Val Gly Ile Val Ser Trp Gly
            165             170             175

Ser Tyr Gly Cys Ala Arg Gly Asn Lys Pro Gly Val Tyr Thr Arg Val
            180             185             190

Ser Ser Tyr Leu Asp Trp Ile
        195
```

```
<210>  5
<211>  43
<212>  PRT
<213>  Homo sapiens
```

```
<400>  5
```

```
Ser Thr Cys Gly Gly Pro Asp Glu Phe Gln Cys Gly Ser Gly Arg Arg
1               5               10              15

Cys Ile Pro Arg Ser Trp Val Cys Asp Gly Asp Pro Asp Cys Glu Asp
            20              25              30

Gly Ser Asp Glu Ser Leu Glu Asn Cys Ala Ala
        35              40
```

```
<210>  6
<211>  2532
<212>  DNA
<213>  Homo sapiens
```

```
<400>  6
atggacaaag aaaacagcga tgtttcagcc gcacctgctg acctgaaaat atccaatatc      60

tcagtccaag tggtcagtgc ccaaaagaag ctgccagtga gacgaccacc gttgccaggg     120

agacgactac cattgccagg aagacgacca ccacaaagac ccattggcaa agccaaaccc     180

aagaagcaat ccaagaaaaa agttcccttt tggaatgtac aaaataaaat cattctcttc     240

acagtatttt tattcatcct agcagtcata gcctggacac ttctgtggct gtatatcagt     300

aagacagaaa gcaaagatgc tttttacttt gctgggatgt ttcgcatcac caacatcgag     360

tttcttcccg aataccgaca aaaggagtcc agggaatttc tttcagtgtc acggactgtg     420

cagcaagtga taaacctggt ttatacaaca tctgccttct ccaaatttta tgagcagtct     480

gttgttgcag atgtcagcag caacaacaaa ggcggcctcc ttgtccactt ttggattgtt     540

tttgtcatgc cacgtgccaa aggccacatc ttctgtgaag actgtgttgc cgccatcttg     600

aaggactcca tccagacaag catcataaac cggacctctg tggggagctt gcagggactg     660

gctgtggaca tggactctgt ggtactaaat gctgggcttc ggtcagatta ctcgtcaacc     720

ataggatctg acaaaggctg ctctcagtac ttctatgcag agcatctgtc tctccactac     780
```

```
ccgctggaga tttctgcagc ctcagggagg ctgatgtgtc acttcaagct ggtggccata      840

gtgggctacc tgattcgtct ctcaatcaag tccatccaaa tcgaagccga caactgtgtc      900

actgactccc tgaccatttta cgactccctt ttgcccatcc ggagcagcat cttgtacaga      960

atttgtgaac ccacaagaac attaatgtca tttgtttcta caaataatct catgttggtg     1020

acatttaagt ctcctcatat acggaggctc tcaggaatcc gggcatattt tgaggtcatt     1080

ccagaacaaa agtgtgaaaa cacagtgttg gtcaaagaca tcactggctt tgaagggaaa     1140

atttcaagcc catattaccc gagctactat cctccaaaat gcaagtgtac ctggaaattt     1200

cagacttctc tatcaactct tggcatagca ctgaaattct ataactattc aataaccaag     1260

aagagtatga aaggctgtga gcatggatgg tgggaaatta atgagcacat gtactgtggc     1320

tcctacatgg atcatcagac aattttttcga gtgcccagcc ctctggttca cattcagctc     1380

cagtgcagtt caaggctttc agacaagcca cttttggcag aatatggcag ttacaacatc     1440

agtcaaccct gccctgttgg atcttttaga tgctcctccg gtttatgtgt ccctcaggcc     1500

cagcgttgtg atggagtaaa tgactgcttt gatgaaagtg atgaactgtt ttgcgtgagc     1560

cctcaacctg cctgcaatac cagctccttc aggcagcatg ccctctcat ctgtgatggc      1620

ttcagggact gtgagaatgg ccgggatgag caaaactgca ctcaaagtat tccatgcaac     1680

aacagaactt ttaagtgtgg caatgatatt tgctttagga aacaaaatgc aaaatgtgat     1740

gggacagtgg attgtccaga tggaagtgat gaagaaggct gcacctgcag caggagttcc     1800

tccgccttc accgcatcat cggaggcaca gacaccctgg aggggggttg gccgtggcag      1860

gtcagcctcc actttgttgg atctgcctac tgtggtgcct cagtcatctc cagggagtgg     1920

cttctttctg cagcccactg ttttcatgga aacaggctgt cagatcccac accatggact     1980

gcacacctcg ggatgtatgt tcaggggaat gccaagtttg tctccccggt gagaagaatt     2040

gtggtccacg agtactataa cagtcagact tttgattatg atattgcttt gctacagctc     2100

agtattgcct ggcctgagac cctgaaacag ctcattcagc caatatgcat tcctcccact     2160

ggtcagagag ttcgcagtgg ggagaagtgc tgggtaactg ctggggggcg aagacacgaa     2220

gcagataata aaggctccct cgttctgcag caagcggagg tagagctcat tgatcaaacg     2280

ctctgtgttt ccacctacgg gatcatcact tctcggatgc tctgtgcagg cataatgtca     2340

ggcaagagag atgcctgcaa aggagattcg ggtggacctt tatcttgtcg aagaaaaagt     2400

gatggaaaat ggattttgac tggcattgtt agctggggac atggaagtgg acgaccaaac     2460

tttcctggtg tttacacaag ggtgtcaaac tttgttccct ggattcataa atatgtccct     2520

tctcttttgt aa                                                         2532
```

<210> 7
<211> 843

9

<212>   PRT
<213>   Homo sapiens

<400>   7

Met Asp Lys Glu Asn Ser Asp Val Ser Ala Ala Pro Ala Asp Leu Lys
1   5   10   15

Ile Ser Asn Ile Ser Val Gln Val Val Ser Ala Gln Lys Lys Leu Pro
            20   25   30

Val Arg Arg Pro Pro Leu Pro Gly Arg Arg Leu Pro Leu Pro Gly Arg
        35   40   45

Arg Pro Pro Gln Arg Pro Ile Gly Lys Ala Lys Pro Lys Lys Gln Ser
        50   55   60

Lys Lys Lys Val Pro Phe Trp Asn Val Gln Asn Lys Ile Ile Leu Phe
65   70   75   80

Thr Val Phe Leu Phe Ile Leu Ala Val Ile Ala Trp Thr Leu Leu Trp
                85   90   95

Leu Tyr Ile Ser Lys Thr Glu Ser Lys Asp Ala Phe Tyr Phe Ala Gly
                100   105   110

Met Phe Arg Ile Thr Asn Ile Glu Phe Leu Pro Glu Tyr Arg Gln Lys
            115   120   125

Glu Ser Arg Glu Phe Leu Ser Val Ser Arg Thr Val Gln Gln Val Ile
        130   135   140

Asn Leu Val Tyr Thr Thr Ser Ala Phe Ser Lys Phe Tyr Glu Gln Ser
145   150   155   160

Val Val Ala Asp Val Ser Ser Asn Asn Lys Gly Gly Leu Leu Val His
            165   170   175

Phe Trp Ile Val Phe Val Met Pro Arg Ala Lys Gly His Ile Phe Cys
            180   185   190

Glu Asp Cys Val Ala Ala Ile Leu Lys Asp Ser Ile Gln Thr Ser Ile
        195   200   205

Ile Asn Arg Thr Ser Val Gly Ser Leu Gln Gly Leu Ala Val Asp Met
    210   215   220

Asp Ser Val Val Leu Asn Ala Gly Leu Arg Ser Asp Tyr Ser Ser Thr
225   230   235   240

10

```
Ile Gly Ser Asp Lys Gly Cys Ser Gln Tyr Phe Tyr Ala Glu His Leu
              245             250             255

Ser Leu His Tyr Pro Leu Glu Ile Ser Ala Ala Ser Gly Arg Leu Met
              260             265             270

Cys His Phe Lys Leu Val Ala Ile Val Gly Tyr Leu Ile Arg Leu Ser
              275             280             285

Ile Lys Ser Ile Gln Ile Glu Ala Asp Asn Cys Val Thr Asp Ser Leu
        290             295             300

Thr Ile Tyr Asp Ser Leu Leu Pro Ile Arg Ser Ser Ile Leu Tyr Arg
305             310             315             320

Ile Cys Glu Pro Thr Arg Thr Leu Met Ser Phe Val Ser Thr Asn Asn
              325             330             335

Leu Met Leu Val Thr Phe Lys Ser Pro His Ile Arg Arg Leu Ser Gly
              340             345             350

Ile Arg Ala Tyr Phe Glu Val Ile Pro Glu Gln Lys Cys Glu Asn Thr
              355             360             365

Val Leu Val Lys Asp Ile Thr Gly Phe Glu Gly Lys Ile Ser Ser Pro
        370             375             380

Tyr Tyr Pro Ser Tyr Tyr Pro Pro Lys Cys Lys Cys Thr Trp Lys Phe
385             390             395             400

Gln Thr Ser Leu Ser Thr Leu Gly Ile Ala Leu Lys Phe Tyr Asn Tyr
              405             410             415

Ser Ile Thr Lys Lys Ser Met Lys Gly Cys Glu His Gly Trp Trp Glu
              420             425             430

Ile Asn Glu His Met Tyr Cys Gly Ser Tyr Met Asp His Gln Thr Ile
              435             440             445

Phe Arg Val Pro Ser Pro Leu Val His Ile Gln Leu Gln Cys Ser Ser
        450             455             460

Arg Leu Ser Asp Lys Pro Leu Leu Ala Glu Tyr Gly Ser Tyr Asn Ile
465             470             475             480

Ser Gln Pro Cys Pro Val Gly Ser Phe Arg Cys Ser Ser Gly Leu Cys
```

11

```
                    485                      490                      495


     Val Pro Gln Ala Gln Arg Cys Asp Gly Val Asn Asp Cys Phe Asp Glu
             500                 505                 510


     Ser Asp Glu Leu Phe Cys Val Ser Pro Gln Pro Ala Cys Asn Thr Ser
             515                 520                 525


     Ser Phe Arg Gln His Gly Pro Leu Ile Cys Asp Gly Phe Arg Asp Cys
             530                 535                 540


     Glu Asn Gly Arg Asp Glu Gln Asn Cys Thr Gln Ser Ile Pro Cys Asn
     545                 550                 555                 560


     Asn Arg Thr Phe Lys Cys Gly Asn Asp Ile Cys Phe Arg Lys Gln Asn
                     565                 570                 575


     Ala Lys Cys Asp Gly Thr Val Asp Cys Pro Asp Gly Ser Asp Glu Glu
             580                 585                 590


     Gly Cys Thr Cys Ser Arg Ser Ser Ser Ala Leu His Arg Ile Ile Gly
             595                 600                 605


     Gly Thr Asp Thr Leu Glu Gly Gly Trp Pro Trp Gln Val Ser Leu His
     610                 615                 620


     Phe Val Gly Ser Ala Tyr Cys Gly Ala Ser Val Ile Ser Arg Glu Trp
     625                 630                 635                 640


     Leu Leu Ser Ala Ala His Cys Phe His Gly Asn Arg Leu Ser Asp Pro
                     645                 650                 655


     Thr Pro Trp Thr Ala His Leu Gly Met Tyr Val Gln Gly Asn Ala Lys
                     660                 665                 670


     Phe Val Ser Pro Val Arg Arg Ile Val Val His Glu Tyr Tyr Asn Ser
             675                 680                 685


     Gln Thr Phe Asp Tyr Asp Ile Ala Leu Leu Gln Leu Ser Ile Ala Trp
             690                 695                 700


     Pro Glu Thr Leu Lys Gln Leu Ile Gln Pro Ile Cys Ile Pro Pro Thr
     705                 710                 715                 720


     Gly Gln Arg Val Arg Ser Gly Glu Lys Cys Trp Val Thr Gly Trp Gly
                     725                 730                 735
```

12

Arg Arg His Glu Ala Asp Asn Lys Gly Ser Leu Val Leu Gln Gln Ala
            740   745   750

Glu Val Glu Leu Ile Asp Gln Thr Leu Cys Val Ser Thr Tyr Gly Ile
            755   760   765

Ile Thr Ser Arg Met Leu Cys Ala Gly Ile Met Ser Gly Lys Arg Asp
            770   775   780

Ala Cys Lys Gly Asp Ser Gly Gly Pro Leu Ser Cys Arg Arg Lys Ser
785   790   795   800

Asp Gly Lys Trp Ile Leu Thr Gly Ile Val Ser Trp Gly His Gly Ser
            805   810   815

Gly Arg Pro Asn Phe Pro Gly Val Tyr Thr Arg Val Ser Asn Phe Val
            820   825   830

Pro Trp Ile His Lys Tyr Val Pro Ser Leu Leu
            835   840

(51) International Patent Classification⁷: C12N 15/57, 9/64, 15/63, 5/10, C12P 21/00, C12Q 1/68, A61K 38/48

(21) International Application Number: PCT/EP01/08993

(22) International Filing Date: 3 August 2001 (03.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/223,479   7 August 2000 (07.08.2000)   US
60/280,132   2 April 2001 (02.04.2001)   US
60/296,440   8 June 2001 (08.06.2001)   US

(71) Applicant (for all designated States except US): BAYER AKTIENGESELLSCHAFT [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and
(75) Inventor/Applicant (for US only): XIAO, Yonghong [US/US]; 78 Dana Street #1, Cambridge, MA 02138 (US).

(74) Common Representative: BAYER AKTIENGESELLSCHAFT; 51368 Leverkusen (DE).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:
25 July 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: REGULATION OF HUMAN MEMBRANE-TYPE SERINE PROTEASE

(57) Abstract: Reagents which regulate human membrane-type serine protease activity and reagents which bind to human membrane-type serine protease gene products can be used to regulate extracellular matrix degradation. Such regulation is particularly useful for treating cancer (e.g., suppressing metastasis of malignant cells), neurodegenerative diseases, osteoporosis, and chronic obstructive pulmonary disease.

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7   C12N15/57   C12N9/64   C12N15/63   C12N5/10   C12P21/00
         C12Q1/68   A61K38/48

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7   C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

SEQUENCE SEARCH, EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 887 414 A (SMITHKLINE BEECHAM PLC) 30 December 1998 (1998-12-30) abstract; claims	1-13, 18-67
A	WO 99 49055 A (BRUCK CLAUDINE ELVIRE MARIE ;SMITHKLINE BEECHAM BIOLOG (BE); COCHE) 30 September 1999 (1999-09-30) abstract; claims	1-13, 18-67
E	WO 01 98468 A (ELLIOTT VICKI S;INCYTE GENOMICS INC ; YUE HENRY (US)) 27 December 2001 (2001-12-27) Seq Id Nos 17, 38 abstract; claims; table 2	1-13, 18-67

-/--

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 May 2002	06/06/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Ceder, O

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
E	WO 02 00860 A (SUGEN INC ;WHYTE DAVID (US); CAENEPEEL SEAN (US); CHARYDCZAK GLEN) 3 January 2002 (2002-01-03) abstract; claims; figures 1QQ,1RR,2Q page 1-6 page 159, line 9 - line 24 page 188, line 1 - line 19 -----	1-13, 18-67

FURTHER INFORMATION CONTINUED FROM   PCT/ISA/ 210

Continuation of Box I.1

Although claims 68-71 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.1

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy

---

Continuation of Box I.2

Claims Nos.:   14-17 68-71

Present claims 14-17 68-71 relate to a reagent defined by reference to a desirable characteristic or property, namely that it modulates a function of a human membrane-type serine protease.

The claims cover all reagents having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such reagents. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the reagent by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search has been carried out for these claims.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
EP 0887414	A	30-12-1998	CA	2231015 A1	09-12-1998
			EP	0887414 A2	30-12-1998
			JP	11075868 A	23-03-1999
			US	6100059 A	08-08-2000
WO 9949055	A	30-09-1999	AU	3519299 A	18-10-1999
			CA	2323760 A1	30-09-1999
			WO	9949055 A1	30-09-1999
			EP	1064388 A1	03-01-2001
			JP	2002507425 T	12-03-2002
WO 0198468	A	27-12-2001	AU	6983601 A	02-01-2002
			WO	0198468 A2	27-12-2001
WO 0200860	A	03-01-2002	AU	7142701 A	08-01-2002
			WO	0200860 A2	03-01-2002